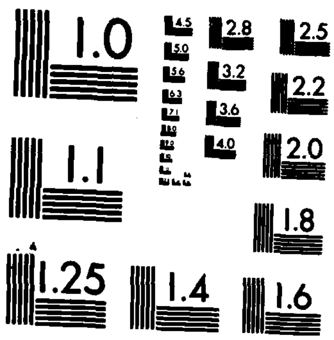
AD-A159 055 PERFORMANCE MEASUREMENT GUIDELINES FOR RESEARCH(U) 1/3
VREULS RESEARCH CORP THOUSAND OAKS CA D VREULS ET AL.
13 JUN 85 AFOSR-TR-85-0642 F49620-80-C-0058

UNCLASSIFIED F/G 5/9 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 85-0642

4

PERFORMANCE MEASUREMENT

GUIDELINES FOR RESEARCH

Vreuls Research Corporation

4

FINAL TECHNICAL REPORT
F49620
13 June 1985

PERFORMANCE MEASUREMENT

GUIDELINES FOR RESEARCH

By:

Vreuls Research Corporation
68 Long Court, Suite E
Thousand Oaks, CA 91360-6084

Donald Vreuls
Richard W. Obermayer
Alan L. Wooldridge
Michael J. Kelly

Contract F49620-80-C-0058

The views and conclusions contained in this document are those of the authors and should not necessarily be interpreted as representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research of the United States Government.

Controlling Office:

Air Force Office of Scientific Research/NL
Bolling AFB, D.C. 20332-6448

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 85-0642	2. GOVT ACCESSION NO. AD-A159055	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Performance Measurement Guidelines for Research		5. TYPE OF REPORT & PERIOD COVERED Final Report 15 Mar 1980 - 30 Sep 1982
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Donald Vreuls, Michael J. Kelly, Richard W. Obermayer, Alan L. Wooldridge		8. CONTRACT OR GRANT NUMBER(s) F49620-80-C-0058
9. PERFORMING ORGANIZATION NAME AND ADDRESS Vreuls Research Corporation, 68 Long Court, Suite E, Thousand Oaks, CA 91360-6084		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2313/A2 61102F
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research, Life Sciences Directorate, Bolling AFB, D.C.		12. REPORT DATE 13 June 1985
		13. NUMBER OF PAGES 204
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for Public Release; Distribution Unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

Sponsored in part by the Naval Training Equipment Center, Orlando, Florida, and the Army Research Institute for the Behavioral and Social Sciences, Alexandria, Virginai.

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Performance Measurement
Segmentation Logic
Transformations

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

All three military services are developing automated human performance measurement systems for aviation training devices and research on human performance. Each organization tends to follow its own course, building on what has been done in the past. If there is no institutional memory, an analysis of what and how to measure starts anew, usually with a literature search; but, performance measurement knowledge is embedded in hundreds of reports, and the measurement details needed to construct measures usually are not found in technical reports which emphasize research issues.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

BLOCK 20. ABSTRACT (Continued)

The purpose of this study was to create a set of aircrew-system performance measurement guidelines for research based on a review of current practice, and the measurement experience and technical judgment of the investigators. The guidelines assume a reader who has introductory knowledge of aviation tasks and aircraft, specific knowledge of the tasks to be investigated, and the services of an instrumentation or simulation engineer to implement the suggested measurement methods. We assume a reader would use this document as a "kit of tools" for guidance on measurement methods and philosophy, and would modify the techniques as appropriate for the specific research task and problem.

A subjective analysis of common measurement requirements among flight tasks for all phases of military aviation was conducted. The selection of system state variables would be dictated by the individual research problem, but guidelines for sampling, measure segmentation, and selection of transforms to create measurement were developed for common flight tasks and measurement problems. Performance measurement issues in system design, training, and automated performance measurement system design were discussed. FORTRAN program listings for common transforms and specialized multivariate data analyses for selecting and constructing measurement from empirical data were appended.

Use of the illustrated techniques was recommended, as was the need to update these techniques as measurement experience accrues. Since measurement is information for a particular purpose, future efforts of this type should analyze user information requirements as well as measurement requirements.

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SECTION 1.

INTRODUCTION

PURPOSE OF STUDY

All three military services are developing automated human performance measurement systems as a part of aviation training devices and for research on training and human performance. Each research and training development program tends to follow its own course, building on what has been done by each investigator or organization in the past. If no institutional memory exists, an analysis of what and how to measure starts anew, usually with a survey of the literature.

The literature can provide examples of measurement and lists of system states that have been recorded (cf. Mixon and Moroney 1982); however, measurement knowledge is embedded in hundreds of technical reports, and the details needed by a researcher or measurement analyst seldom are provided. Also, many tasks for research are so constrained that generalization of measurement may be tenuous. Simply, the "book" on flight performance measurement has not been written.

An initial goal of this research was to search for a common core of measurement information which would be needed across many flight tasks and environments, to provide a basis for partial standardization of measurement practice, at least for research. But because measurement is information for a specific purpose, it became evident during the course of the study that such a lofty goal might not be achieved without a thorough analysis of all flight maneuvers and tasks, along with an identification and analysis of all research purposes and information needs. The level of effort which would be required to perform an analysis of this magnitude was well beyond what could be done within the scope of this study.

Recognizing the difficulties, the investigators were asked to use the available information and data, known measurement practices, and their own experience and technical judgment. The purpose of the study, then, became one of describing common aircrew measurement problems and solutions in the form of guidelines.

PURPOSE OF REPORT

The purpose of this report is to document general measurement issues and propose solutions, where possible, recognizing that measurement for all research cannot be addressed. There are useful measurement tools and practices that have evolved over the years, and recent breakthroughs which have improved real-time performance measurement systems. This report is intended to provide a framework for measurement, and suggest state-of-the-art approaches. It should be updated as new information and experience with performance measurement becomes available.

At the onset of the study the investigators were advised to focus on measurement, not assessment, using the following distinction: Measurement

is information about performance. Assessment requires the use of many sources of information to determine the quality of performance for a particular purpose, such as the goodness or badness of performance relative to criteria for training or operations. This clearly was beyond the intent of the effort, although there is a need for operational figures of merit and criteria, as discussed in Section 8. Also, if measurement is viewed as information for a specific purpose, the distinction blurs somewhat.

ORGANIZATION OF REPORT

This report addresses the measurement of performance of human aircrew members as controllers of an aircraft system and its subsystems. It assumes the reader has an introductory knowledge of aircraft and flight tasks, specific knowledge of the tasks that are to be investigated, and would use this document as a "kit of tools" for guidance on measurement methods. It also assumes the user has the services of a simulation or instrumentation engineer to mechanize the rules and algorithms presented here within the appropriate reference and computational systems that are to be used.

Section 2 discusses measurement commonality among the various maneuvers and tasks for all phases of flight. It shows that algorithms for measuring turns, climbs and descents, accelerations and decelerations would provide a general measurement capability for tasks performed in nearly one-half of all phases of military aviation.

Section 3 defines the structure used in this report for describing measurement in terms of system states, sampling rates, measure segments, desired values, error data and transformations. Typical system state variables (often referred to as "parameters") of interest are discussed in Section 4, along with guidelines for sampling.

Measure segmentation rules are discussed in Section 5. Segmentation of maneuvers, such as aerobatic and basic fighter maneuvers, is so dependent on the aircraft and the desired measures that these tasks are discussed separately in Appendix A. Various transformations and the information they provide are discussed in Section 6, and FORTRAN subroutines for generating common transforms are listed in Appendix B.

An important use of measurement is to diagnose performance for research and training. Measurement system intelligence and data requirements for performance diagnosis are discussed in Section 7. Comments on statistical data analysis for measurement development purposes are offered in Appendix C, along with FORTRAN programs. These data analysis methods are unique; they provide methods for selecting measures from a set of candidates and overall metrics for use in training and research. This description and documentation does not appear in any other document in its present form.

Section 8 reviews the flight performance measurement domain, and current issues in measurement for system design, training, and research. Whereas earlier sections of the report treat what we know and have learned recently about measurement, this section discusses what we need to learn to advance the state-of-the-art. Our conclusions and recommendations are in Section 9.

SECTION 2.

COMMONALITY ANALYSIS

The purpose of the study was to seek common aircrew performance measurement problems and suggest measurement guidelines where possible. Across the domain of flight, different types of aircraft fly similar missions and maneuvers, and many of the same tasks are performed over and over again for various maneuvers. For example, both fixed wing and helicopter pilots have to navigate from point-to-point. They have to (a) capture a designated track across the ground (or over water), (b) maintain that track within prescribed limits, (c) arrive at navigational fixes according to flight plan time estimates, and (d) depart fixes to capture and maintain a new track, heading or course. Although the accuracy of performance, ease of workload and specific control techniques for doing the above navigational functions may depend on the method and system being used for navigation, the functional tasks repeat over and over again.

This example illustrates that at a functional level, there is a common core of tasks. The common core of functional tasks may not lead to a complete, "necessary and sufficient" set of measures for any particular task, but it should lead to those measures of performance and measurement processing considerations which an analyst would need to consider in almost all cases.

An analysis was performed to find the level of task description that might guide the selection of maneuvers and flight tasks to be used as a basis for common measurement guidelines. A fully documented, hierarchical task analysis was well beyond the scope of the effort; rather, a judgmental review of flight missions and tasks was done, based on the experience of the authors. To find a common core of flight tasks and measurement problems, flight tasks were reviewed at the "mission" level by aircraft type and at the "maneuver" level for all aircraft.

COMMON FLIGHT REGIMES BY AIRCRAFT TYPE

Military aviation can be classified into ten arbitrary flight regimes shown in Table 1. These regimes represent a mixture of training stages and missions for fixed wing, rotary wing and V/STOL aircraft. Flight regimes are not mutually exclusive categories; there are common tasks across the regimes. Transition includes basic flight tasks, take-off and landing. Utility and transport missions have been omitted from the list because it is unnecessary to separate them from transition, navigation and instrument regimes from a measurement viewpoint. Also, carrier qualification (CQ) is considered to be a special case of transition training for the Navy, just as Nap of the Earth (NOE) flight for Army rotary wing operations is considered as a special case of navigation from a measurement viewpoint.

Table 1 shows that V/STOL and fixed wing aircraft fly the same flight regimes except airdrop. Within the common flight regimes, V/STOL and fixed wing aircraft are inseparable except during (a) vertical takeoff, hover and

landing, (b) transitions to and from forward aerodynamic flight, and (c) vectoring in forward flight during air combat.

TABLE 1. FLIGHT REGIMES BY AIRCRAFT TYPE

Flight Regime	Aircraft Type: Fixed Wing	Rotary Wing	V/STOL
1. Transition	*	*	*
2. Navigation	*	*	*
3. Formation	*	*	*
4. Instruments	*	*	*
5. Aerobatics	*		*
6. Basic Fighter Maneuvers	*		*
7. Air-to-Air Combat	*		*
8. Air-to-Surface Combat	*	*	*
9. Air Refueling	*		*
10. Air Drop	*		

* Regime commonly flown.

Rotary wing aircraft operate in many of the same flight regimes as fixed wing aircraft. At present, they do not do (a) aerobatics, (b) basic fighter maneuvers, (c) air refueling, (d) air drop or (e) air-to-air combat, although helicopter air-to-air-combat is a possibility in the future. It is likely that future helicopter air-to-air tactics will differ from conventional wing tactics, and will require unique measurement.

Insofar as measurement of functional tasks is concerned, the type of aircraft is of little consequence except in specific regions of unique capability, such as hovering flight and transitions to and from forward aerodynamic flight. The ability to hover, translate and transition to aerodynamic flight affects the measurement of V/STOL and rotary wing aircraft maneuvers such as takeoff and landing (from land bases or ships). It also affects the measurement of rotary wing performance of NOE and air-to-surface combat tasks and weapons delivery. Measurement for these flight tasks and maneuvers should be treated separately.

COMMON MANEUVERS

Flight tasks and maneuvers can be described in many different ways, ranging from Instructional System Development (ISD) statements of terminal learning objectives and performance standards to matters of technique (how to do it) found in training manuals for particular aircraft. If one lists every flight maneuver that has been given a name, several hundred maneuvers result. For example, there are cross-wind landings, short field landings, and no flap landings; in each case the performance objectives are to contact the runway with (a) a minimum rate of descent, (b) minimum lateral speed (drift), (c) minimum error from the runway centerline and (d) enough runway remaining to stop. There are different techniques for landing in different circumstances which might be the subject of diagnostic measures, but one would always measure descent rate, drift, centerline error and distance down the runway for a field landing.

For purposes of maneuver commonality analysis for measurement, general categories of maneuvers were combined; 75 maneuvers were classified into 12 flight phases shown in Figure 1. Figure 1 is a two dimensional matrix which illustrates the commonality between all maneuvers. No attempt was made to quantify the degree of relationship between maneuver pairs; an "x" simply illustrates a "non-trivial" relationship between each maneuver pair based on our experience. The highlights of this analysis are discussed in the following paragraphs:

Basic Flight. Fundamental aircraft maneuvers are to turn, climb and descend, and accelerate and decelerate (or to hold various combinations of their rates at given values, which may include zero as in straight and level flight). The fixed wing pilot controls pitch and roll attitude, yaw, thrust and drag to do these maneuvers. In forward aerodynamic flight, the helicopter is controlled the same way as a fixed wing aircraft, except there is direct lift control with collective, and yaw must be controlled by the pilot to "trim" the aircraft (e.g. point the aircraft in the direction it is traveling).

Turns are divided into "normal" and "steep." In a steep turn the magnitude of the control problem increases dramatically as aircraft lift vector approaches horizontal, and lift, thrust, and g-loading limits are approached; expected performance accuracy changes, and different measures become appropriate. Turns which require about 35 degrees of bank angle or less (20 degrees for helicopters) are considered normal turns. All others are considered steep turns.

For purposes of this analysis, a distinction is made between "air reference" and "ground reference" maneuvers. For example, a 720 degree steep turn at constant altitude, turn rate and speed is considered an air referenced maneuver because the turn and speed is to be maintained relative to the air mass. A constant radius turn around a pylon (on the ground) would be considered as a ground referenced maneuver during which the pilot would have to adjust the turn rate for wind drift. Ground referenced maneuvers may contain both ground and air mass referenced tasks.

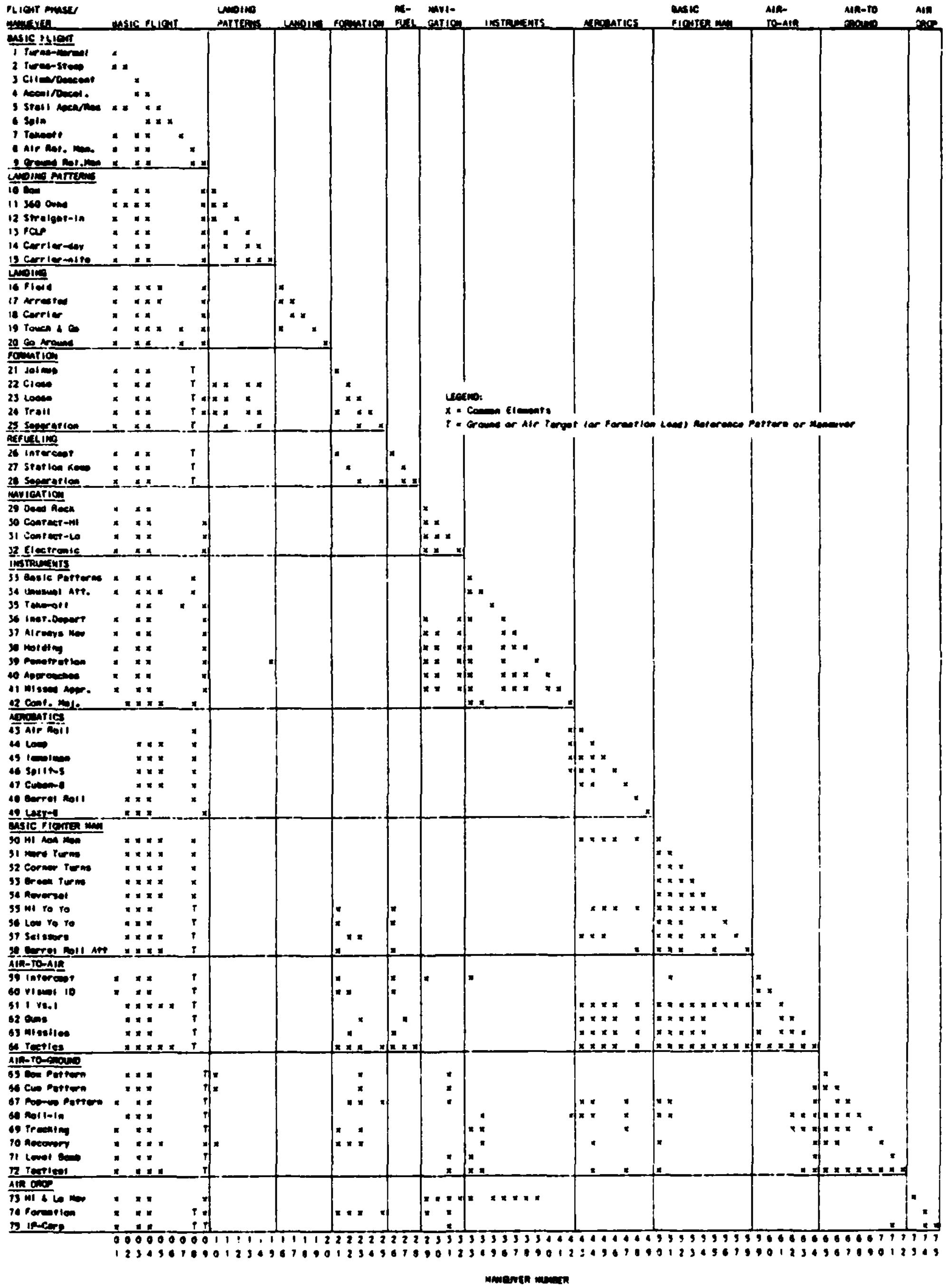


Figure 1. Commonality of Flight Maneuvers for Measurement.

[Parenthetically, one must know earth axis, aircraft body axis, and inertial axis coordinate systems and transformations between them when designing measurement, but it is not necessary to discuss these systems for purposes of commonality analysis.]

Some basic flight maneuvers are special cases of the fundamental maneuvers. The approach to stall and stall recovery are considered as special cases of acceleration and deceleration, where the usual control actions might not have the same effect as when the aircraft is at a lower angle of attack. Stalls and spins are departures from controlled flight; they require special control actions that are airframe dependent. Takeoff is a special case of acceleration (and ground referenced turn--no turn, just hold heading) followed by a climb and acceleration.

It is obvious that the ability of the pilot to turn, climb, descend, accelerate, approach stall and maneuver the aircraft relative to the air mass and ground is fundamental to (and shares much commonality with) the remaining flight phases and tasks shown in Figure 1. The commonality of the remaining flight phases with basic flight maneuvers will not be mentioned in the discussion which follows to avoid redundancy.

Landing Patterns. In general, there are three basic patterns, the box pattern, the 360 degree overhead, and the straight in; they are all variations on the theme. The box pattern has four segments, upwind (usually the take-off runway centerline extension), crosswind (90 degrees from upwind), downwind (180 degrees from upwind), base (90 degrees from final), and final (the runway heading). The 360 degree overhead pattern has five segments, initial (starting about three-to-five miles out on the runway heading but usually displaced from the final approach course slightly), break (a steep turn usually abeam the runway to dissipate energy and turn onto downwind), downwind, base and final. A straight in approach has only one leg, final.

Field Carrier Landing Practice (FCLP) is a special case of the 360 degree overhead approach to simulate day aircraft carrier patterns, where the speeds and altitudes are lower than the normal 360 overhead approach, and precise speed, turn and spacing (from other aircraft) control are emphasized. Day carrier patterns are more like instrument flight maneuvers than normal airport patterns because of the required precision; however, visual reference to the carrier is maintained to intercept the FLOLS (Fresnel Lens Optical Landing System) glideslope reference on the turn to final. Spacing is most important because the prior aircraft in the pattern must be cleared from landing area, and typically, many aircraft have to be recovered in a short time. The night carrier landing pattern is a straight in approach. These patterns are separated from the basic patterns because the precision required during carrier approach is much greater (and suggests slightly different measures) than a normal land-based landing pattern.

Landing. It is judgmental where an approach ends and landing begins. From an instructional viewpoint, approach and landing are one continuous task; however, from a measurement viewpoint, they can be separated, and it is useful to do so to develop general measurement algorithms. For this

analysis, we assume that the approach ends at the threshold of the approach end of the runway, or the ramp on an aircraft carrier, and the landing begins at that point.

There are four classes of landings: field, arrested, carrier and touch and go. The go-around (or missed approach) is treated arbitrarily as a special case of landing (although it could be treated just as well as a special case of takeoff). This reasons for separating these classes of landings are discussed below:

The field landing is the normal landing in which there is a target airspeed to be at over the runway threshold, an intended touchdown point beyond the runway threshold, and a flare maneuver will be done. The normal field landing includes the rollout on the runway, which may require thrust reversal, braking and parachute deployment for deceleration.

Arrested landing (engaging a crash barrier) may take place at the approach or departure ends of the runway, may or may not involve a flare, and may require special procedures which depend on the airframe and the characteristics of the arrestment device used; for these reasons, arrested landings are different from normal field landing for measurement purposes, and have to be treated on a case by case basis.

Carrier landings are precise controlled descents to engage the aircraft tail hook on a target arrestment cable (usually, number 2 of four cables) on the carrier deck. There is no flare maneuver, and little last minute maneuvering is possible. Wind speed and direction over the deck usually is constant, but there are turbulence and downdraft effects because of carrier superstructure. Procedures demand application of full thrust on impact so that power is available if there is no hook engagement (this is called Bolter). There is no rollout.

Touch and go landings usually are similar to normal field landings, but the procedures may vary slightly with particular airframes; for example, full flaps might not be used if a touch and go is contemplated, and the approach airspeed might (or might not) be different. Once the aircraft touches down, takeoff power is applied, and the aircraft is reconfigured as necessary for takeoff. Typically, the aircraft is close to takeoff rotation speed at touchdown, so reconfiguration (flaps, trim and power) has to be done quickly. Measurement for touch and go landings may not be a simple concatenation of field landing and takeoff measures because of the need for aircraft reconfiguration while rolling, and the possibility that rotation airspeed may be exceeded when the reconfiguration is taking place; in short, the measures of importance may change slightly.

For the same reasons, a go-around maneuver is not just a climb task. There are intervening tasks to reconfigure the aircraft to an accelerate and climb configuration (power, flaps, gear and trim) while minimizing altitude loss but maintaining safe airspeeds. Usually, a go-around is caused by a misjudged approach or conflicting traffic; avoiding collision with ground obstacles and the other aircraft are important elements of the task. If the aircraft is at a minimum fuel state, the go-around procedures

may vary slightly from a normal takeoff and climb, necessitating different measurement.

Although performance of the landing task is dependent on a proper landing pattern and approach, the landing maneuvers share few common elements with landing pattern tasks (but share many common elements with basic flight tasks).

Formation. There are three general tasks in formation flight: joinup, stationkeeping and separation. The joinup is an intercept problem, where the pilot has to maneuver the aircraft into an assigned position about the lead aircraft, which also is moving through the air. Stationkeeping is maintaining position even though the lead aircraft may maneuver; changes in position, such as cross-overs or cross-unders are viewed as special cases of stationkeeping for this analysis. For purposes of measurement, there are three types of stationkeeping tasks: close, loose and trail.

The purpose of the close formation is to move several aircraft as one unit through the airspace. The number two, three or four aircraft must maintain an exact position about the lead aircraft.

The purpose of the loose formation has more to do with tactics and lookout doctrine than moving several aircraft in one unit of airspace; here, an approximate position about the lead aircraft is to be maintained in a more fluid fashion for ground reconnaissance and visual surveillance of areas the other aircraft cannot see well.

The trail formation is used in some low level flights; here, the objective is to fly a precise distance behind the lead aircraft, and fly the exact track over the ground as the lead aircraft (e.g. turning where it turned, and not at the same time it turned).

Separation from a formation may take several forms; usually, there is a planned procedure for each aircraft to leave the formation at a specific time (or the occurence of a specific event), and fly a predetermined flight profile to insure safe separation of aircraft.

Formation flight shares common elements with landing patterns because visual discriminations and the control of aircraft relative to the movement of other aircraft are common tasks.

Refueling. Refueling is similar to formation flight, with the additional requirement of maintaining contact with the tanker refueling element while flying close to the tanker. The task is difficult because the "formation" position places the aircraft to be refueled in a dangerous position relative to the wingtip vortex, turbulence, and slipstream aerodynamic effects of the tanker. Also, the weight and balance of both aircraft are changing during fuel transfer; the stability and control characteristics of the aircraft being refueled change dramatically, and fuel transfer and management may impose additional tasks on the pilot or flight crew. Although there are many common elements with formation flight, there are additional refueling tasks that require measurement.

Navigation. There appear to be four kinds of navigation tasks: dead reckoning, contact high, contact low and electronic. The fundamental navigation method is dead reckoning--flying a course at a given ground speed for a given time should move an aircraft from a present position to a predictable new position with certain accuracy. Of course, winds aloft change, heading may vary, gyros precess, magnetic variation changes, and over long distances, the curvature of the earth's surface is a factor. Contact navigation and electronic systems are used to correct the inaccuracies of dead reckoning.

Contact navigation is the use of visual reference to the terrain and maps to determine present position. High and low altitude contact navigation differ slightly. High altitude courses usually are straight lines or rumb lines, are flown at altitudes which are clear of obstacles, and would be performed by reference to predominant terrain features, such as large bodies of water, islands, whole mountain ranges, major rivers, and urban areas. Low altitude navigation courses may be composed of straight lines for brief intervals, curved courses (to follow valleys, canyons or river beds), are flown at altitudes which are not clear of obstacles (and for Nap-of-the-Earth (NOE) flight, at or below the tree tops), and are flown by reference to terrain details such as hilltops, river branches, bridges and culverts, road intersections, and cultural details. High and low altitude contact navigation differ in workload (frequency of required position determinations, number of checkpoints per unit time, and the required attention to terrain and obstacle avoidance maneuvering). The measurement techniques for high and low altitude navigation may vary, although some common measures such as checkpoint accuracy and flight plan time accuracy are common.

Dead reckoning navigation can be corrected (or replaced) by electronic systems. These systems include radio navigation aids such as automatic direction finding (ADF), visual omni-directional radio range facilities (VOR), distance measuring equipment (DME), TACAN, radar terrain mapping, doppler radar and inertial navigation systems. Measures of aircrew performance using these systems may shift in emphasis from instrument interpretation and aircraft control performance to data entry and system management accuracy as the level of automation increases. Data entry and system operation measures may be needed in addition to course errors, checkpoint accuracy and flight plan time accuracy if the cause of a navigation error is to be determined; but, this measurement will be system specific.

There is commonality between navigation tasks, as illustrated in Figure 1, but little commonality with the maneuvers which have been discussed above. Navigation is a fundamental task which is required in most of the flight phases in the remainder of Figure 1.

Instruments. Instrument flight shares a high degree of commonality with basic flight and navigation tasks. The fundamental difference is that all maneuvers are done by reference to instruments and electronic systems alone. Ten types of maneuvers are listed in Figure 1 and discussed briefly below.

Basic patterns are simply air referenced profiles which are used to practice the fundamental maneuvers (turns, climbs, descents, speed and altitude changes) for precise aircraft control. Unusual attitude recoveries are for practicing recognition of impending emergency (such as stall or exceeding maximum airspeed) situations, and executing a safe recovery with a minimum altitude loss and without overstressing the airframe.

Measurement of an instrument takeoff is similar to measuring a visual takeoff. Instrument departures, airways navigation, penetrations, approaches and missed approaches are measured in the same way that navigation performance is measured; there is a high degree of commonality between these maneuvers and navigation tasks.

The holding pattern task requires unique measurement treatment if diagnostic performance information is desired. Fix crossing, turns, outbound drift correction, inbound turns and drift correction, and inbound course intercept and tracking all have to be measured in a minimum amount of time with few data samples. The implementation of automated diagnostic holding pattern measurement has not been achieved to date.

The final instrument maneuvers are the confidence maneuvers. These are aerobatic maneuvers such as the aileron roll, loop, immelman, chandelle and split-s, which are used to practice interpretation of instruments in all flight attitudes and promote pilot confidence. These maneuvers share some common elements with unusual attitude recoveries, and are nearly identical to aerobatic maneuvers which follow.

<u>Aerobatics</u>. Fundamentally, aerobatic maneuvers introduce the concept of turning, maneuvering, and managing aircraft potential and kinetic energy in the vertical plane. Except for instrument confidence maneuvers, vertical plane maneuvering and energy management in the previous phases of flight has been limited to climbing and descending. There are several air show type aerobatic maneuvers (such as the hammer-head stall, eight point roll, inside snap roll, outside snap roll, outside turn, and square loop) that are uncommon in military aviation. Figure 1 lists seven common military aerobatic maneuvers.

The aileron roll requires the pilot to roll the aircraft about its own longitudinal axis without changing heading, whereas the barrel roll is a roll about an axis which is displaced laterally from the aircraft, and requires pitch, roll and heading changes. The loop is a 360 degree turn in the vertical plane (loops which start with a pitch-up are inside loops). The immelman is a half inside loop, or 180 degree turn which uses only the vertical plane; the aircraft does a half roll near the top to return from inverted to normal flight. The split-s is a 180 degree turn in the vertical plane which is entered by rolling inverted, then completed by performing the last half of an inside loop.

The cuban eight is a horizontal figure eight which is similar to an immelman which terminates in a 30-45 degree dive (instead of the top of the loop) which is followed by another immelman terminating in a 30-45 degree dive, and so forth. A chandelle (not listed in Figure 1) is a maximum per-

formance climbing turn which uses both lateral and vertical planes; it is a combination of a steep climbing turn and an immelman.

The lazy eight is a pitch, roll and speed coordination task which is ground referenced. It requires simultaneous climbing and descending turns so that a horizontal figure eight is described about a selected reference point on the horizon. The maneuver starts from straight and level flight with wings pointed at a reference point on the horizon. A 90 degree climbing turn is started, but when 45 degrees have been turned, pitch is decreased so that it is 0 degrees at 90 degrees of turn (the top of the climb); during this time, bank angle is increased to nearly 90 degrees at the top of the climbing turn. At this point, a 90 degree descending turn is continued, but when another 45 degrees have been turned, pitch is increased so that it is 0 degrees at 180 degrees of turn (the bottom of descent); during this time, bank angle is decreased so that the wings are level and point to the horizon reference at the bottom of descent. There are specific airspeed targets at the top of climb and bottom of descent. The second half of the maneuver is similar to the first half, but in the opposite direction.

Figure 1 shows there are common elements between the aileron roll and the immelman, split-s and cuban eight, which share common elements with the loop (but the loop and aileron roll have no commonality). The barrel roll and lazy eight share little in common with the other aerobatic maneuvers, except for energy management and vertical plane maneuvering. The first 45 degrees of the lazy eight is similar to a chandelle entry, and elements of this maneuver are similar to nose-high unusual attitude recovery on instruments. Using the vertical plane to turn and manage energy are essential components in basic fighter maneuvers, air-to-air combat and air-to-ground weapons delivery.

Basic Fighter Maneuvers. The classic aerial dog fight involves two aircraft, each of which is trying to maneuver into a position within weapons effectiveness range, and hold that relative position long enough to fire weapons with enough accuracy to destroy the opponent. Modern weapon systems and tactics may make the classic one-versus-one dog fight a rare event in combat, but the fundamental skills remain important for survival. If the pilot does not have an all-aspect missile system (and except for a lucky head-on shot, or high angle raking gunshot), the dog fight is a game to see who can turn the tightest (or otherwise maneuver) to position himself behind the opponent long enough to achieve a firing solution.

Basic Fighter Maneuvers (BFM) are a prerequisite to air-to-air and air-to-ground combat maneuvers. A sample of BFM listed in Figure 1 fall into three categories, high angle-of-attack maneuvering, maximum performance turns and maneuvering relative to an opponent (who is also maneuvering).

High angle-of-attack maneuvering is the foundation task; pilots must learn to control their aircraft when turning, climbing or diving under high g-loads at near stall conditions. Techniques will vary with the airframe, but in general, excessive use of aileron for bank control is hazardous, and swift but smooth control of pitch is mandatory to prevent a departure. In

some aircraft, the stick forces to generate increasing amounts of g-loads DECREASE as g-increases; small amounts of stick pressure and movement will change angle-of-attack significantly.

There are three classes of maximum performance turns. A hard turn is the maximum turn rate which can be sustained without loosing airspeed at a given altitude and gross weight at maximum thrust. A corner turn is the maximum turn rate which can be achieved at a given gross weight and altitude. Typically, there is only one airspeed which will yield a corner turn (at a given gross weight and altitude); attempts to turn tighter will either produce a stall or exceed airframe g-limits. A corner turn cannot be held for long because typically there is more drag generated than thrust available to overcome drag. A break turn is the maximum turn rate that can be achieved at any airspeed; if the maximum instantaneous turn rate is held, speed will decay until a corner turn is achieved.

In addition to maximum performance turns, there are five types of maneuvers which are used to gain advantage of the opponent. A reversal is switching almost instantaneously from a hard turn in one direction, to a hard turn in the opposite direction, typically to escape from a position of disadvantage (e.g. the opponent is about to shoot you down). The high and low Yo-Yo's use the vertical plane (and gravity) to tighten the turn about an opponent who is turning in the horizontal plane. The scissors involve pitching and rolling maneuvers (vertical or horizontal) to get behind the opponent without sacrificing more energy than necessary. A barrel roll can use the vertical plane to reduce lateral separation. A barrel roll attack uses the vertical plane to turn and position one aircraft above and behind a second aircraft, who's track is crossing ahead of the first aircraft.

Figure 1 shows there are common elements between BFM and aerobatics (as would be expected). Within BFM, maneuvers share many common elements, which can be summarized as high angle-of-attack maneuvering, maximum performance turning, and the use of the vertical plane to change direction and slow down relative to an opponent while sacrificing as little airspeed as possible. BFM introduces the need to measure performance relative to the movement and intention of another aircraft which is not cooperative; also, high angle of attack maneuvering and maximum performance turns introduce measurement sampling and algorithm challenges which are not found in other flight phases. Aerobatics and BFM share common elements with the following air-to-air and air-to-ground combat maneuvers.

<u>Air-to-Air</u>. A sample of air-to-air combat maneuvers is illustrated by six types of maneuvers listed in Figure 1; these are intercept, visual identification, one-versus-one, guns, missiles and tactics. The intercept and visual identification maneuvers share common tasks with formation and refueling maneuvers; but, they may be performed using onboard electronic or external systems, and the target is not cooperative. Moreover, detection by the opponent must be avoided, which makes the task more complex than formation or refueling maneuvers; tactical considerations, such as the capability and intent of the adversary and the rules of engagement, enter the task.

One-versus-one combat is the application of BFM. Fuel consumption, energy management and energy maneuverability share equal importance with BFM maneuvering tasks. Often, fuel is limited, so any potential engagement will be constrained by fuel. Kinetic energy (speed) and potential energy (altitude and fuel) must be traded-off for advantage when maneuvering to obtain positional advantage. Energy maneuverability is the task of maintaining the engagement within the envelope of speed, turn rate and energy gain or loss which is more advantageous for the aircraft you are flying than the same factors for the opponent aircraft; this is critical when engaging an aircraft of dissimilar performance capability. (For example, if you are in an F-4 engaging a MIG-21, you do not want to slow down to the MIG's corner turn speed unless a kill is assured, because the MIG will be able to turn tighter at that speed than you can.) These tasks add considerable perceptual and cognitive workload to BFM tasks.

A complementary, but separate set of skills are required for gun and missile delivery tasks. There are weapon system management tasks of selecting, configuring, arming and firing the desired weapons; these tasks have to be executed quickly and at the proper time during maneuvering, when everything else is going on. The aircraft must be controlled to achieve a firing solution for gun and missile delivery tasks without losing a position of advantage. For guns, the target has to be tracked in the gunsight for a period of time as well. These tasks share common elements with formation, refueling, aerobatics and BFM tasks, but add gunsight and radar scope tracking tasks.

Typically, military air-to-air combat is not planned to be one-versus-one dog fights. Tactics for two-versus-one, two-versus-two and so forth are taught and practiced. Although the execution of such tactics would include all the previously described tasks, there are added tasks of mission planning, contingency planning, inflight coordination, command, control and communications. Pilots must use their aircraft as an element of a coordinated strike team.

Air-to-Ground. Eight maneuvers shown in Figure 1 provide a sample of air-to-ground weapons delivery maneuvers. The first three maneuvers are typical patterns. The next three are common tasks in manual dive bombing, strafing and rocket delivery. The last two maneuvers represent level bomb and tactical deliveries.

Three different kinds of delivery patterns are used in training to position the aircraft at the proper point in the sky to roll-in to a dive (and to control aircraft separation during training). The box pattern is similar to a landing pattern, except that the downwind leg altitude is a function of the dive angle, and can vary from 10,500 feet to about 5,000 feet. A high or low angle dive is started on the turn from base to final leg.

The cone pattern is flown by intercepting and flying just outside of an imaginary inverted cone which has an apex above the target and sides which represent the dive angle path to the target. The cone is intercepted on the downwind leg abeam the target, and flown for about 90 degrees at pattern altitude before rolling-in to the final dive.

The pop-up pattern is a tactical maneuver used to minimize exposure of the aircraft to target air defense systems. Typically, the maneuver will be entered from high speed, low level flight at a distance from the target which is just enough to complete the maneuver. A nominal climb (about 30 degrees of pitch-up) is established. Just before the apex altitude (which depends on the final dive angle), the aircraft is rolled inverted and pitched into the final dive, making any required changes in course to line-up with the target. As the desired dive angle is achieved, the aircraft is rolled upright to complete the delivery.

The above patterns are used to fly the aircraft to the proper point in the sky from which a delivery of known dive angle can be made. All three patterns share common maneuvers of roll-in, target tracking and recovery. The roll-in from a box or cone pattern typically requires a bank angle of 90 degrees plus the dive angle (e.g. a 30 degree dive requires a 120 degree bank). The "roll-in" equivalent in the pop-up pattern is apex. As the proper dive angle and sight picture is achieved, the aircraft is rolled upright, the weapon system is armed, and target tracking begins.

Since the aircraft has a lift vector, the dive angle path will be less than the dive angle attitude; the sight "pipper" is usually placed between the aircraft and the target, and is permitted to drift to the target as the dive continues. The tracking maneuver is to control aircraft attitude so that required airspeed, dive angle and pipper placement are achieved simultaneously with the release altitude. Adjustments to pipper placement and release altitude have to be made for off parameter solutions. All these tasks have to be accomplished in three to eight seconds, depending on dive angle and entry altitude. For strafing, the target has to be tracked when firing the guns or cannon.

After release or cease firing of guns, a four-g recovery maneuver is executed so as not to stall or strike the ground. Weapons systems are switched to safe. A pull-up and climbing turn is executed to position the aircraft with proper spacing behind the prior aircraft when using the box or cone patterns. In a tactical delivery, the aircraft probably would return immediately to low level flight to escape the target area.

The final two classes of maneuvers shown in Figure 1 are level bomb and tactical deliveries. Level bomb shares common navigation to target and system configuration tasks with the above maneuvers, but no dive is made; the aircraft flies at constant altitude (high or low) to the target, and releases its bombs when the proper sight picture is achieved.

There are several types of tactical deliveries which depend on the weapon systems available on specific aircraft. Radar Nav-Bomb involves the use of radar returns for navigation to the target and for determining weapons release aim points; these deliveries are similar to level bomb. Modern weapons systems with head-up displays can compute and display bomb fall lines and projected impact points; a predetermined dive angle does not have to be achieved, but the range of delivery angles may be constrained by bomb fragmentation patterns, target terrain or tactical considerations; the most typical maneuver would be the pop-up or a variation of it.

Modern weapon systems permit delivery maneuvers which "toss" the bomb at the target as well. These maneuvers require designation of the target to the weapon system, then following commands which are usually presented on a head-up display. Typically, these maneuvers start with a pitch-up, then command release at a certain point in the pitch-up maneuver. If the bomb is to be tossed to a target which is ahead of the aircraft, the release will occur during the early part of the pitch-up. If an "over the shoulder" delivery is used against a target which is below the aircraft, the release will occur late in the pitch-up maneuver, in a near vertical climb. Depending on the pitch attitude at release, the pitch-up is converted to a steep descending turn or continued to an immelman to escape the target area.

Figure 1 shows that air-to-ground weapons delivery maneuvers share common tasks with landing patterns, formation flight, low altitude contact navigation, instrument flight, aerobatics, basic fighter maneuvers, and air-to-air combat. There are many common tasks among the air-to-ground weapons delivery maneuvers as well.

Air Drop. The typical airdrop mission is flown in formation and under radio silence. After take-off, the formation will assemble and may fly to a location using airways, but eventually will descend and fly low level contact navigation in extended trail formation to an Initial Point (IP). At the IP, there is a slow-down maneuver which allows the elements of a typical three ship formation to close on either side of the lead for the delivery formation.

Between the IP and the Computed Air Release Point (CARP), pilots must maintain their formation position while the lead navigator recomputes the CARP based on current winds, and loadmasters prepare for delivery. The drop starts at the CARP, which may be determined with the aid of onboard avionics, but is usually a computed number of seconds from passing abeam a visual landmark near the drop zone. Formation aircraft drop when the lead drops, and pilots must adjust for changing weight and balance of the aircraft when holding formation position. After delivery the loadmasters close exits, the formation executes a planned acceleration maneuver, and returns (typically) to a low level flight in trail formation for a return to base or other designated airport.

Airdrop shares common tasks with formation, navigation, instruments, and air-to-ground maneuvers.

Summary. The number of common flight maneuvers among flight phases is summarized in Figure 2. The number of x's in each off-diagonal cell of flight phases in Figure 1 were divided by the total number of cells formed by each phase pair to derive the percentage of cells filled shown in Figure 2. Percentages less than 10 were judged to be beyond the accuracy of this analysis, and were not shown.

Flight Phase	BAS	L-P	LND	FOR	REF	NAV	INS	ARO	BFM	A-A
BAS Basic Flight										
L-P Landing Patt.	.46*									
LND Landing	.56									
FOR Formation	.49	.43								
REF Refueling	.44			.27						
NAV Navigation	.42									
INS Instruments	.47					.43				
ARO Aerobatics	.40									
BFM Basic Fighter	.53			.11	.11			.21		
A-A Air-to-Air	.52			.30	.39			.43	.54	
A-G Air-to-Ground	.47			.25		.16		.18		.25
A-D Air Drop	.52			.27		.58	.20			

(Column heading: Flight Phase)

* Number of common maneuvers divided by total per phase pair.

Figure 2. Commonality of flight maneuvers by flight phase.

Accepting the subjective foundation of this analysis, Figure 2 shows commonality between the maneuvers in each pair of flight phases. Basic flight maneuvers share common maneuvers with all phases, as expected. The relationship between landing patterns and formation flight is because of the task of maneuvering aircraft about other aircraft in traffic patterns; it may be overestimated. Landing shares no commonality with other flight maneuvers.

Formation flight has common elements with refueling, basic fighter maneuvers, air-to-air, air-to-ground, and air drop. Refueling has common elements with basic fighter maneuvers and air-to-air. Navigation shares common elements with instruments, air-to-ground and air drop. Instruments and air drop are related only because of possible airways navigation during the air drop mission. Aerobatics share common elements with basic fighter maneuvers, air-to-air and air-to-ground. Basic fighter maneuvers share common tasks with air-to-air. And, air-to-air shares some common tasks with air-to-ground.

One must be cautious with a quantitative analysis of the data in Figure 2; the relationships are between flight phase pairs, one pair at a time, and there are no corrections for multiple relationships. A qualitative analysis is possible; it suggests that measurement of basic flight tasks would account for about half of the measurement for any of these flight phases. Re-examining Figure 1, it would be useful to develop general guidelines for measuring turns, climbs and descents, accelerations and decelerations, stall approach and recovery, air and ground referenced maneuvers.

Measures for formation intercept and stationkeeping (close and loose) appear to have utility for basic fighter maneuvers, air-to-air, air-to-ground and air drop. Measures of navigation performance appear to have utility for instruments, air-to-ground and air drop. Aerobatics measures should be useful for basic fighter maneuvers, air-to-air and air-to-ground. Basic fighter maneuver measures should aid air-to-air measurement. And, air-to-air measurement should aid air-to-ground measurement development.

This analysis has not examined flight planning, subsystem operations (such as fuel, hydraulic, electrical, and avionics subsystems), emergency procedures, electronic warfare and command, control and communications tasks. Development of performance measures for these tasks are (a) system and situation specific, (b) may be beyond the scope of what is reasonable to measure with automated systems or (c) are beyond the scope of what could be accomplished in this study. With the commonality analysis results in mind, a structure for common measurement is discussed next.

SECTION 3.

MEASUREMENT STRUCTURE

It is useful to think about measurement as deriving information for some purpose. The general purpose of measuring aircrew performance is to quantify performance for system design, personnel selection, training, operational readiness evaluation and research for these purposes. What the measurement user wants to know will dictate what measures are appropriate. Often, however, measurement users do not know what they want to know with much precision; this is especially true in research, which by definition is a quest for previously unknown knowledge. To improve the precision of measurement specification, it is helpful to define a structure, or language for describing measures.

It is assumed that a measurement user needs performance metrics which are amenable to quantitative analysis. Plots of variables over time (a time history), although useful for measurement development and performance diagnosis, do not yield metrics which are suitable for quantitative analysis. The process of generating performance metrics can be decomposed into the following functional elements:

- Measure Segment.
- State Variable(s).
- Sampling Rate.
- Desired Value (if any).
- Error Datum.
- Transformation.

Unambiguous definition of these elements permits data on the states of the system or human controller to be converted into a unit of information about performance. Each element is introduced in the remainder of this section, and described at length in remaining sections of this report.

MEASURE SEGMENT

A measure segment is any period of time during which the aircrew/system performance is lawful, and for which the beginning and end can be defined unambiguously. Measure segments generally correspond to tasks. They can overlap, and several segments can be in operation at a given time. They can represent a whole maneuver, or component parts of a maneuver, depending on what the measurement user wants to know.

For example, a takeoff can be divided into many segments such as brake release to rotation speed, rotation speed to lift-off, and lift-off to gear retraction for measuring acceleration and pitch control performance. At the same time, a segment for the whole maneuver can start at brake release and end at gear retraction for measuring lateral displacement from the centerline throughout the whole maneuver.

As will be discussed in Section 5, precise and unambiguous logic for determining when to measure is a key issue of real-time measurement system design.

STATE VARIABLE(S)

A state variable is any measurable state of the human or system, such as control inputs, aircraft attitudes, rates, heading, speed, altitude and physiological states of the human operator. In many documents, including simulator specifications, state variables often are called parameters, but this use of the term is incorrect. A parameter is: "A quantity to which the operator may assign arbitrary values, as distinguished from a variable, which can assume only those values that the form of the function makes possible" (Webster, 1956). What has been called parameters are variables, which can be continuous or discrete. Common state variables are listed in Section 4.

SAMPLING RATE

State variables are sampled for measurement purposes. Typically, the sampling rate is expressed as the number of samples per unit of time, such as ten samples per second. The sampling rate depends on how fast the variable changes value per unit of time and the accuracy required for the transformation (see definition below) to be used. Common sampling rates are discussed in Section 4.

DESIRED VALUE AND ERROR DATUM

Often, the information required by the measurement system user is the magnitude of the error between a present value of a variable and its ideal or desired value at that time. For example, if a pilot is supposed to fly at 250 knots and the current airspeed is 262 knots, the error from the desired speed is 12 knots, which is the error datum. A desired value may be a constant or a function of other variables.

TRANSFORMATION

A transformation is any logical or mathematical treatment of the error datum. A measurement user may want to know only the error of a particular continuous variable at a given time, such as the error from a desired speed when crossing the runway threshold. Here, the transformation would be the actual (or absolute) value of the error datum. Continuous variables often are sampled for a period of time to describe performance with one metric; here, a transformation such as the absolute average error over the measure segment might be used. Transformations may be performed on one variable (or error data), or may be formed by functions of several variables or error data. Transformations may be in the time or frequency domains, as discussed further in Section 6.

INTERACTIONS BETWEEN MEASUREMENT STRUCTURE ELEMENTS

The functional elements of the measurement structure are dependent on one another; there are interactions between segmentation rules, desired values and transforms. If measure segmentation is activated on one set of system state variables, then typically, resultant measures will be made on other variables.

For example, during instrument climbing and descending turns, there are specific altitude changes to be achieved at certain headings which should occur at specific times during the maneuver. The maneuver may be segmented by time, altitude or heading. If segments are activated by heading, then altitude and time errors when achieving specific headings become measures. If segments are activated when achieving certain altitudes, then heading and time errors become measures.

Also, it is obvious that maneuver segmentation rules and sampling rates place boundaries on desired values and transformations that can be used. For example, if a maneuver segment is simply one sample to be taken when an event occurs (e.g. pitch attitude, airspeed and altitude at gear retraction after takeoff), transformations are limited to the value or absolute value of the error data, or functions of these three measures. If the frequency content of pilot control input is desired, the sample rate will have to be high enough for accurate measurement of the frequency spectrum.

These interactions illustrate that the meaning or information content of measures depends on an unambiguous specification of all elements of this measurement structure. Most crew/system performance measurement can be defined using this structure, as discussed further in remaining sections of this report.

SECTION 4.

TYPICAL SYSTEM STATES AND SAMPLING RATES

As discussed in Section 3, system state variables are measurable states of the human or system, such as control inputs, aircraft attitudes and rates, heading, speed, altitude, position and acceleration about other aircraft, subsystem controls and displays, communications and physiological states of the human operator. Physiological states of the human operator can include temperature, heart rate, respiration rate, oxygen intake, eye movements and so forth, but these are beyond the intent of this report, and will not be discussed.

The selection of system states will depend entirely on the purpose of measuring. If one is measuring instrument flight performance, a relatively small number of states, perhaps 20-30, will be needed. If one is measuring performance for combat effectiveness and pilot workload assessment during system design studies, several hundred states of the pilot, all subsystems, adversary systems and weapons effectiveness will be needed.

As a guide, typical system states that one might examine for a basic and advanced single engine jet training simulator are shown in Table 2, along with the accuracy required for human performance measurement. A nominal set of subsystems (such as a standard UHF and VHF communications and navigation, a manual gunsight, two hydraulic systems, and master caution alerting systems) are assumed. The table also shows typical aircraft states relative to other aircraft, ground or air targets for formation, manual air-to-air and manual air-to-ground weapons delivery tasks. A multi-engine jet fighter, large bomber or transport aircraft would have many more systems than are shown for fuel, engine, hydraulic, electrical, navigation and weapons system control.

Table 2 shows typical sampling rates for human performance studies. The required sampling rate depends on how fast the variables change value per unit of time and the accuracy required for the transformation to be used. For example, variables such as course deviation may not change rapidly; when using statistical transformations, one sample per second may be more accurate than is necessary. Statistical transformations such as average or absolute average deviations, root-mean-squared error and the standard deviation are robust; if the measurement segment is one minute or more in length and there is not much maneuvering (as in instrument flight), these transformations will represent performance accurately with sampling rates as low as one sample every 10 seconds.

On the other hand, in tasks such as air-to-air and air-to-ground weapon delivery, variables change rapidly, and the highest possible sampling rate is required. Also, when sampling pilot control inputs for frequency domain transformations, 20 samples per second may be required. One rule of thumb is to sample at a rate which is at least three times the highest expected frequency of the variable, although five times the highest frequency is safer. Transformations are discussed at length in Section 6.

TABLE 2. TYPICAL SYSTEM STATES AND SAMPLING RATES

System States	Accuracy	Sample Rate	Type
Pitch Stick Position or Force	1% of range	20/sec	C*
Roll Stick Position or Force	1% of range	20/sec	C
Rudder Pedal Position or Force	1% of range	20/sec	C
Throttle Position	1% of range	2/sec	D
Collective Position (helicopter)	1% of range	2/sec	C
Speedbrake Switch Position	actual	2/sec	D
Flap Selector Position	actual	2/sec	D
Hook Position (carrier aircraft)	actual	2/sec	D
Pitch Trim (Position)	0.1 deg	2/sec	D
Aileron Trim (Position)	1.0 deg	2/sec	D
Rudder Trim (Position)	1.0 deg	2/sec	D
Altimeter Setting	0.01 in	2/sec	D
Pitch Attitude	1.0 deg	2/sec	C
Pitch Attitude Rate	0.1 deg/sec	2/sec	C
Roll Attitude	1.0 deg	2/sec	C
Roll Attitude Rate	0.1 deg/sec	2/sec	C
Heading (magnetic)	1.0 deg	2/sec	C
Angle of Attack	0.5 units	2/sec	C
Airspeed	2.0 knots	2/sec	C
Mach	0.01 mach	2/sec	C
Turn Rate (needle position)	0.25 needle	2/sec	C
g's (indicated)	0.5 g	2/sec	C
Side Slip (ball position)	0.1 ball	2/sec	C
Altitude (barometric indicated) .	10.0 feet	2/sec	C
Vertical Speed (climb/dive indicated)	50.0 ft/min	2/sec	C
RPM (indicated)	1%	2/sec	C
TIT or EGT (indicated)	10.0 deg	2/sec	C
Engine Start Switch	actual	2/sec	D
Ignition button	actual	2/sec	D
Start/relight button	actual	2/sec	D
Fuel Selectors	actual	2/sec	D
Central Warning Panel Lights	actual	2/sec	D
Fire Extinguisher Pushbutton and Light	actual	2/sec	D
Hydraulic Pressure (each system)	1.0 bar	2/sec	D
Parking brake lever	actual	2/sec	D
Anti-Skid switch	actual	2/sec	D
Brake Parachute Test switch	actual	2/sec	D
Brake Parachute Control	actual	2/sec	D
Brake Hydraulic Pressure (each system)	1.0 bar	2/sec	D
Landing Gear Controls (up, down, emgcy)	actual	2/sec	D
UHF Communications (Freq/Mode/X-mit)	actual	2/sec	D
Transponder Mode/Code/Ident	actual	2/sec	D
TACAN T/R switch, Frequency	actual	2/sec	D

* C = Continuous, D = Discrete

TABLE 2. TYPICAL SYSTEM STATES AND SAMPLE RATES (Continued)

System States	Accuracy	Sample Rate	Type
TACAN Radial	1.0 deg	2/sec	C
HSI Course Deviation (indicated)	1.0 deg	2/sec	C
TACAN DME	0.1 nm	2/sec	C
DME Lock (DME Validitiy)	actual	2/sec	D
DME Rate	0.1 nm/hour	2/sec	C
VOR Frequency, Omni Bearing Selector	actual	2/sec	D
ADF Frequency	actual	2/sec	D
ADF Bearing	1.0 deg	2/sec	C
ILS Localizer Deviation	0.01 deg	2/sec	C
ILS (or FLOLS or ACLS) Glideslope	0.01 deg	2/sec	C
Aircraft Latitude (or feet from Ref)	5.0 sec	2/sec	C
Aircraft Longitude (or feet from Ref)	5.0 sec	2/sec	C
Magnetic Track across ground	1.0 deg	2/sec	C
Weight off nose gear	actual	2/sec	D
Weight off main gear (both)	actual	2/sec	D
Fuel Remaining	10.0 pounds	2/sec	C
Gross Weight	100.0 pounds	2/sec	C
Stores Loaded (jettesoned)	actual	2/sec	D
Anti-Ice	actual	2/sec	D
Pitot Heat	actual	2/sec	D
Additional Key Cockpit Switches	actual	2/sec	D
Add For Formation:			
Delta X from A/C 1**	1.0 foot	10/sec	C
Delta Y from A/C 1	1.0 foot	10/sec	C
Delta Z from A/C 1	1.0 foot	10/sec	C
Delta X, Y, Z from A/C 2	1.0 foot	10/sec	C
Delta X, Y, Z from A/C 3	1.0 foot	10/sec	C
Add For Air-to-Ground:			
Target to Pipper (in gunsite plane)	1.0 mil	20sec	C
Gunsight Mil Setting and Controls	actual	1/sec	D
Trigger Pull	actual	20/sec	D
Weapon Control Panel (select/arm/displays)	actual	1/sec	D
Bomb Impact Circular Error	1.0 foot	1/sec	D
Bomb Impact Direction	5.0 degrees	1/sec	D
Gun Rounds Fired	actual	1/sec	D
Number of Hits	actual	1/sec	D
Add For Air-to-Air:			
In range	100.0 feet	10/sec	D
In envelope	5.0 deg	10/sec	D
Missles fired	actual	10/sec	D
Missle kill	actual	10/sec	D
Line of Sight Angle (cone)	5.0 deg	10/sec	C
Target Aspect Angle	5.0 deg	10/sec	C
Delta X, Y, and Z Target Aircraft	1.0 foot	10/sec	C
Delta X, Y, and Z Rate of Target A/C	20.0 ft/sec	10/sec	C

** Delta is the difference in X, Y, and Z between aircraft.

Typically, it is convenient, but not necessary to measure all states of the aircraft. For example, attitude rates can be derived from attitude changes over time if necessary, but it is better to measure a rate or acceleration directly than to derive it by digital methods. Similarly, turn rate can be derived from the change in heading over time; however, as the turn approaches the vertical plane in maneuvering flight, body axis or inertial axis turn rate is needed. A case-by-case analysis of information desired, sensors and states available for measuring is required.

Sampling of discrete variables such as trigger for weapons release, switch positions, frequencies, and data entry into weapons and flight management systems requires special treatment. Some switches can change to a critical state and return to the previous state (or another state) rapidly. It is possible to miss a state unless the switch is sampled at the highest possible rate of change that can occur.

This is obvious during weapons delivery tasks, but not so obvious when measuring normal procedures of subsystem control. If one examines the number of switches in a typical jet aircraft, and the number of possible positions of those switches, an enormous number of measures can be taken; this tends to clog the data storage medium (memory, magnetic disk or tape) with thousands of samples of measures, only a few of which may be useful.

There are three solutions to this problem. First, localized routines which sample at the highest possible rate can be used to remember any switch position changes that occur between the normal measurement sampling rate for data treatment or storage. For example, switches may be examined for position 30 times a second, and all changes reported to the measurement system for further treatment once a second.

Second, sampling of switches can be done whenever any switch changes position, rather than at a fixed time interval (also, sampling of slow moving system states can be done on this basis as well). If sampling is done when states change, a time stamp has to accompany the record so that the "time history" can be reconstructed for measurement purposes.

Third, variable sampling rates can be used. Here, sampling rates of various states can be altered as the situation dictates. Records of the sampling rates in effect have to be maintained, and algorithms to treat each sampling rate appropriately have to be developed.

The best solution to this problem will be situation dependent. In general, however, measurement system design and software is simplified if a consistent sampling method is used for all variables. The first solution, localized routines to remember changes in state between general sampling intervals, is preferred. The second solution is preferred if there are long periods of time, few changes in the relevant states, and transformations do not involve statistical or frequency domain treatment. The third method is preferred when the first method cannot be implemented and there are storage limitations.

This concludes our discussion of system states and sampling rates. Some example measures for a typical tracking task are contained in Section 6, Transformation Guidelines, after treatment of measure segmentation, which is next.

SECTION 5.

MEASURE SEGMENTATION GUIDELINES

As discussed in Section 3, a measure segment is any period of time during which the aircrew-system performance is lawful, and for which the beginning and end can be defined unambiguously. Measure segments generally correspond to tasks. They can overlap, and several segments can be in operation at any given time. They can represent a whole maneuver, or component parts of a maneuver, depending on what the measurement user wants to know. Measure segmentation rules generally apply where there are known profiles or prescribed procedures; however, the concepts can be applied to aid measurement in emergent situations, such as air combat maneuvering, where there may not be any one profile.

Measure segmentation rules influence the precision of measures taken within a segment. Most of flight can be described as transitions to and from steady states. For example, if the task is to turn to a heading, the pilot will bank to capture and track a steady-state roll angle, then roll-out on the heading. If error from the desired steady-state roll angle is measured throughout the turn, more error variance can accrue during the capture of that steady-state than during the tracking of it; this would contaminate tracking measures with capture performance, and decrease the precision of measurement. Also, when the pilot deliberately departs from the steady state (such as rolling out of the turn), that event must be detected to prevent the tracking measures from being contaminated by the departure from the steady-state, which is a transition to the new steady-state.

Measure segmentation rules, therefore, are important because they affect the precision of measures taken within a segment. They also interact with and define the meaning of measures, as discussed earlier. This section will address general issues, general principles, maneuvers and tasks of interest, required logic decisions, example logic and combining logics.

GENERAL ISSUES

Flight Segments or Tasks. The first general issue is: is measurement to be derived for flight segments or pilot tasks within and between flight segments? For example, an instrument flight task might be to fly an airway from intersection A to VOR B, and then to VOR C, where a 45 degree change in course is required from A-B to B-C. Flight segments would be A-B and B-C. The pilot task, however, is to track the course A-B, and at some point just prior to crossing B (for jets), initiate a turn to depart A-B so as to roll-out on B-C on course. The flight tasks are track A-B, transition to B-C, then track B-C. There are three pilot tasks, but only two flight segments.

In slower, non-jet aircraft and helicopters, it is good practice to cross a fix before turning to intercept the outbound leg. In this case, the transition occurs after the fix crossing, but there still are three pilot tasks and two flight segments.

Recommendation: Measure operator tasks, not flight segments, unless they coincide.

Amount of Segmentation. The second general issue is: how much task segmentation is necessary to describe a given performance? In the turn example given above, it was assumed that the measurement user would want to separate the capture and departure of roll angle from the steady-state tracking of roll. This kind of information might be desired for some applications, but not for others. Enough information about performance might be derived from simply timing the turn, from beginning to end, to determine if the average turn rate was within the expected range; in this case, the beginning and end of the turn would have to be detected with precision, but the segmentation of capturing and departing the steady-state turn might not be needed.

Recommendation: An analysis of user information needs is critical. Do not subdivide a pilot task or maneuver any more than necessary to obtain desired performance information (and nothing more) with precision. When a measurement system with good segmentation capacity is developed, there may be strong urges to decompose all tasks and maneuvers to the same level of detail. Do not decompose tasks unless the information is necessary.

Detect Transitions. The third general issue is to develop measure segmentation logic which can detect transitions. Transitions to and from steady-states represent much of flight, yet most measurement examines only the steady-state. Anecdotal evidence suggests that instructors learn much about the skill of students by how they control transitions, and certainly smooth and efficient control of transitions represents deeper knowledge and skill than maintaining a steady-state, which is simply regulator behavior; much performance information can be lost if transitions are not measured.

Although it is easy to detect the transitions to and from steady-state tracking when examining a time history (e.g. a plot of a state variable over time), measurement system intelligence and memory is required to detect transitions in real time. For example, in real time it is hard to know if a pilot is (a) deliberately entering a turn, (b) simply correcting a course or heading error, or (c) not flying with precision. After a turn has taken place, one can determine from the time history plot where it began. During maneuvering flight, the pilot may transition from one state to another continuously, never settling on one state long enough for it to be considered as a steady-state.

Recommendation: It generally is worthwhile in the long run to design a measurement system with short-term memory of the time history, and enough intelligence to be able to start a segment at some earlier time, based on what has happened over a short period of time. Sufficient intelligence can result from the use of short-term memory, functions composed of several variables, flags to denote when events occur and common boolean operators. These elements are discussed in a later subsection.

Precision and Robustness. The fourth and final general issue is the precision and robustness of measure segmentation algorithms. The purpose

of segmenting is to gain meaning and precision from the resulting measures. The logic must capture the performance it was intended to capture with precision, and nothing else. Also, the logic must perform decisions in exactly the same way, no matter how the maneuver or measure segment is entered. There are many ways, for example, that pilots might capture a course, altitude, heading, speed or position relative to another aircraft or object on the ground. Measure segmentation logic must always work in the same way at all times in spite of the foibles and variability of human operators.

Recommendation: Never assume logic is fool-proof. Install protective logic in the measurement system in case segments fail to start or stop. Pretest all logic with a range of "worst case" situations with a sample of actual data, or a simulation of those data.

GENERAL PRINCIPLES

Since it is easy to know when transitions start by looking backward on a time history, and difficult to make such decisions when only the current states of the system are known, real time measurement systems must have short-term memory. The segmentation logic must be able to start a segment or capture an event at some time in the past, as a function of what has happened over a short time interval and what is happening now. The general principles for designing segmentation logic are embodied in the concepts of (a) windowing, (b) Boolean logical operations, (c) flags to denote certain events (d) derivatives of variables, and (e) protective logic. Each concept is described below:

Windowing. A running "window" of all variables and data of interest is recommended. The purpose of the window is to maintain a short time history for segmentation decisions and measurement; a window of 15-30 seconds of time history should be adequate for most measurement purposes. Windowing was first suggested as a solution to logic problems by Hennessy, Hockenberger, Barneby and Vreuls (1979). The window should contain the following functional elements:

- All variables to be measured during this segment and the next one, and any transformations of them (such as tolerance bands or running averages) which are computed "on the fly."

- All computed functions (of perhaps several variables).

- All variables on which segmentation logic will be based for this segment and the next one.

- All binary flags to indicate when specific events occur.

The window does not need to carry all variables and flags of interest for all possible segments; but if the window is restricted to data which are of current interest to the measurement system, it is mandatory to include all variables and flags which apply to the next segment as well. Because of backward looking logic, a decision to open the next segment can

be made retroactively, and the window must already contain the variables, functions and flags that will apply to that next segment.

The window can be envisioned as a two-dimensional array in computer memory, where required variables, functions and flags represent one element of the array, and successive samples of those variables represent the second dimension. In a current application, 32 words of data sampled two times a second are preserved in the window, yielding a 16 second time history. When the window is filled, the oldest sample is lost to make room for the most recent sample. The number of variables, functions, transformations and flags to be carried in the window depend on user information needs, memory space available, and the time required to process the window. These factors have to be determined for each application; but in general, 20 variables per aircraft appear to be sufficient for most flight tasks. Measurement of cockpit and subsystem switches would add to this.

Flags. For segmentation logic purposes, it is helpful to set flags when certain events occur, then test flags as well as variables to make decisions. For example, a flag can be set as the result of logic to detect when certain conditions occur. Later, these conditions may no longer exist, but it may be necessary to know that they once existed, and at what time they existed. Setting a flag, and testing that flag at a later time provides this intelligence, which is needed for maneuvering flight, and to simplify sequential logic where one would look for a certain set of events only after another event occurred.

Boolean Operators. Common AND, OR and NOT operators are sufficient for most segmentation logic decisions. A typical construction for segmentation decisions would be as follows: IF the error from the desired value of variable X is greater than X' AND the error from desired value of variable Y is less than Y', AND flag N is true, THEN start segment Z. The need for more complex constructions is reduced greatly by the use of flags.

Derivatives. The derivative of the error from the desired value of a variable contains information which is useful for segmentation logic when it is used in conjunction with the error from the variable. For example, instantaneous vertical velocity is the derivative of altitude error. Position error (altitude error from desired) can be compared to vertical velocity (rate of change of position error) at any (or every) moment in time to determine if the pilot is (a) displaced from proper position but converging on it (e.g. correcting), (b) displaced from proper position and diverging away from it, (c) displaced from proper position and holding a steady offset, (d) at proper position but diverging from it, or (e) remaining stable on position.

During typical steady-state tracking, a pilot will vary about a desired value of a system state, constantly nulling error in much the same way as a regulator or thermostat. There will be a small amount of error and error rate that represents acceptable boundaries of control performance and stability; this has been called the "limit cycle" by some control system engineers. If the pilot is slightly off altitude but correcting smoothly, or on altitude but diverging from it slowly, the performance probably would be judged to be stable and within the normal limit cycle of the system.

When the system departs from the normal limit cycle, the sign of the position error and rate offers a quick test of convergence or divergence: If position error and rate are opposite in sign, the pilot is converging on the desired position. If the position error and rate are of the same sign, the pilot is diverging from the proper position. When performance is outside the limit cycle, therefore, a simple test of the sign of error can determine whether the pilot is correcting or departing a given state.

This concept is useful, for example, to distinguish the start of a maneuver from normal corrections of the previous steady-state; however, it is not a sufficient test because the pilot may be momentarily unstable. If a flag is set whenever the position error and derivative indicate divergence, and the presence of that flag over several seconds of time is tested, the decision that a change of state is in progress can be enhanced. The use of a window of the last several seconds of time history permits the segmentation logic to detect divergence from a previous steady-state for a long enough period of time to decide that the change is not a normal correction (or momentary instability), then designate the start of a new segment at an earlier time. The segment start time can be based on a nominal (constant) time, a running average of the rate of change per unit of time, or when the flag of error and error rate first indicated continuous divergence.

Parenthetically, a plot of position error versus error rate is known as a "phase plane" in traditional servomechanism control engineering (circa 1940's). Regions of the phase plane provide control performance and stability information, and can be used for segmentation logic decisions as described above, or as a transformation.

Protective Logic. It is hazardous to assume that all segments will start and stop, because all conditions for so doing might not be met. If a segment does not start, there is little that can be done other than to analyze the failure and build more comprehensive start logic. Once opened, all segments should have parallel logic to close them in case the flight terminates for some reason. Also, it is good practice to estimate the maximum possible time of every segment, and close them on timeout. A stop for any reason other than a normal logic decision should be noted in the measurement system output, and measures invalidated.

MANEUVERS AND TASKS OF INTEREST

As said before, much of flight is transitions to and from steady-state turns, climbs and descents, and accelerations and decelerations to capture turn rates or headings, climb rates or altitudes, airspeeds, courses across the ground, or a position relative to an object on the ground or in the air. As discussed in Section 2, performance of these tasks accounts for nearly one-half of all of flight tasks. General rules for segmenting turns, climbs and descents, accelerations and decelerations, therefore, would represent the most useful set of guidelines. The generality of measure segmentation rules for the following tasks are listed below in descending order of value to all phases of flight:

- Turns.
- Acceleration and Deceleration.
- Climb and Descent.
- Capture of Course, Heading, Speed, Altitude or Altitude Rate.
- Stationkeeping.
- Maneuvering relative to another aircraft or an object on the ground.

REQUIRED LOGIC DECISIONS

Segmentation logic has to determine if the aircraft is capturing, tracking or departing:

- Turn rate or bank angle;
- Climb or descent rate;
- Altitude, heading, course or arc;
- Navigational fix or holding pattern;
- Formation station;
- Aerobatic maneuver segments;
- Basic fighter maneuver segments;
- Air combat maneuver segments or positions of advantage or disadvantage;
- Ground attack maneuver segments; or
- Hover, or translation to and form forward aerodynamic flight.

For the above cases, exact algorithms are required, and there are some arbitrary decisions that have to be made, such as when does:

- Capture stop and tracking start?
- Tracking end and transition begin?
- Approach end and go-around begin?
- Approach end and landing begin?
- Landing end and rollout begin (Weight on wheels? How much weight? One or both wheels? Which bounce?)?
- Rollout end?
- Takeoff begin?

EXAMPLE LOGIC

The development of proper and robust segmentation logic is an inexact science. Philosophy and methods can be illustrated, but application of these methods to every problem without thorough analysis, modification to suit the situation, and empirical test is not recommended. For purposes of illustration only, a high performance, fixed wing jet aircraft is assumed in Table 3, unless otherwise noted.

TABLE 3. EXAMPLE MEASURE SEGMENTATION LOGIC

Segment		Start/Stop Logic	Comment
1. Turn			For half standard rate, standard rate and constant bank angle turns.
	Capture:	When Roll vs. error from initial heading is outside the limit cycle and divergent in the proper direction for 5 seconds, START when divergence first began.	Nominal, starting limit cycle values are 10 degrees of Roll and 5 degrees of heading error. This should be determined for each situation. Slower aircraft change heading faster for a given Roll angle.
	Track:	When Roll within 10 degrees of Target Roll angle and Rollrate within 1 degree per second, START.	Standard rate turn Target Roll = IAS * 0.167. Half standard rate Target Roll = IAS * 0.100.
	Rollout:	When Heading within (Target Roll * 0.2) and Target Roll vs. Rollrate divergent for 5 seconds in proper direction, START Rollout 5 seconds earlier or when divergance first began.	
	Stop:	Heading within 5 degrees of desired and Roll within 5 degrees of wings level for 3 seconds.	More complex logic would be required if rollout is not within 5 degrees.
	Reversal:	When Roll passes wings level, STOP previous turn segment and START new turn.	
2. Climb/Descent			For constant Airspeed or Vertical Velocity (VV) climb or descent schedule.
	Capture:	When Altitude error vs. VV divergent in proper direction for 5 seconds, START 8 seconds earlier or when divergance first began.	The Altitude here is the initial altitude. A nominal limit cycle is 150 feet of Altitude error and 500 feet per minute of VV for jets. Reduce VV limit to 250 fpm for slower aircraft.

TABLE 3. EXAMPLE MEASURE SEGMENTATION LOGIC (Continued)

Segment	Start/Stop Logic	Comment
2. Climb/Descent (Continued)		
Track:	When VV within 100 feet per minute of Target, or Airspeed within 5 knots of Target, and Pitchrate within 0.5 degrees per second, START.	Target VV or Target Airspeed are the desired values for the climb or descent.
Leveloff:	When Altitude remaining is less than 10% of VV, and VV is divergent from Target VV (or Airspeed is divergent from Target Airspeed) for 5 seconds, START levelout 5 seconds earlier, or when divergence first began.	Special case for penetration (jet) descent is to first reduce to one-half of descent rate when 1,000 feet above desired leveloff altitude.
Stop:	When Altitude within 100 feet of desired, and VV within 100 feet per minute for 3 seconds, STOP.	More complex logic may be needed if leveloff within 100 feet does not occur.
Reversal:	When Pitch passes horizon, STOP previous segment and START new one.	
3. Acceleration/Deceleration		For acceleration, deceler-tion, and holding constant speed.
Capture:	When Airspeed error from initial speed greater than 5 knots for 10 seconds in the proper direction, START 10 seconds earlier, or when divergance first occurred.	Note: Power should increase for climbing or level speed increase. Power should decrease for climbing or level speed decrease. Power might not increase for descending speed increase.
Track:	When Airspeed within 5 knots of Target Airspeed, and Airspeed Rate within 1 knot per second, START.	Pitch Rate might replace Airspeed Rate, or in some cases rate terms may not be needed. Use Track logic also for reversal.
Depart:	Track segment initiated; use Capture logic, substituting Target Airspeed for initial speed.	

TABLE 3. EXAMPLE MEASURE SEGMENTATION LOGIC (Continued)

Segment		Start/Stop Logic	Comment
4.	Altitude		For steady-state altitude holding.
	Capture:	Use Climb/Descent Leveloff logic.	
	Track:	Use Climb/Descent Stop logic.	
	Depart:	Use Climb/Descent Capture logic.	
5.	Heading		For steady-state heading holding.
	Capture:	Use Turn Rollout logic.	
	Track:	Use Turn Stop logic.	
	Depart:	Use Turn Capture logic.	
6.	Course		For capture, track and depart VOR/TACAN Radial or ADF Bearing.
	Capture:	When Radial (or Bearing) error converging (e.g. the signs of Radial error and Radial error rate are opposite), START.	
	Track:	When the absolute value of Track Crossing error and the absolute value of Radial error are less than 4 degrees (5 degrees for Bearings), for 5 seconds, START.	Track Crossing error is the angle between the desired radial (over the ground) and the projected ground track of the aircraft at each instant in time. Heading can be used if there is no wind.
	Depart:	When Track is active and absolute Radial error is greater than 4 degrees (5 degrees for Bearings) and absolute Track Crossing error is greater than 8 degrees and Radial error is diverging (e.g. the signs of Radial error and Radial error Rate are the same), then START when Radial error first went divergent (or the limits of the window, whichever is least).	Note: The problem with radial or bearing departure is that the rates of change can be affected by distance from the station, and to a lesser extent by whether one is inbound to, or outbound from the station. This general logic might be simplified for specific situations by using either Track Crossing Angle error or Radial error rate, but not both.

TABLE 3. EXAMPLE MEASURE SEGMENTATION LOGIC (Continued)

Segment		Start/Stop Logic	Comment
7.	Ground Track		For capture, track and depart of a track (magnetic or true) across the ground, rumb line, great circle or oceanic.
	Capture:	When Cross Track error and cross track error rate show convergance, START.	Cross Track error is the distance from the track along a line perpendicular to it, and may be in miles, nautical miles, feet, yards or meters.
	Track:	When Capture is TRUE and Cross Track error and error rate are within the limit cycle for 15 seconds, START 10 seconds earlier.	The limit cycle will vary. For low level fixed wing flight a course width of 4 miles is common, and error rate for the limit cycle should be less than 0.1 miles per minute. For helicopter Nap of the Earth (NOE) flight much higher error rates can be expected; we recommend a limit cycle of 0.3 times the acceptable course width.
	Depart:	When Track is TRUE and Cross Track error and error rate are outside the limit cycle and divergent for 15 seconds, START 15 seconds earlier.	Note: This algorithm has never been tested with data. It is reasonable for fixed wing aircraft, but NOE flight may need a longer time test.
8.	DME Arc		For capture, track and depart of a DME or TACAN Arc,or an Arc from a know point on the ground.
	Capture:	When a previous segment has ended and DME error and error rate are converging in the proper direction, START capture.	Typically, the capture of an Arc will occur after crossing a prior fix, or transitioning from a previous segment.
	Track:	When Capture segment is TRUE and DME error and error rate within the limit cycle for for 15 seconds, START 10 seconds earlier.	Limit cycles will vary with Arc radius and airspeed. A nominal limit cycle is 0.5 mile and 0.1 mile per minute, which should be adjusted with trial data.

TABLE 3. EXAMPLE MEASURE SEGMENTATION LOGIC (Continued)

Segment	Start/Stop Logic	Comment
8. DME Arc (Continued		
Depart:	When Track is TRUE and DME error and error rate exceed the limit cycle and are divergent in the proper direction for 15 seconds, START 15 seconds earlier.	
9. Navigation Fix		
Crossing:	Use the Closest Point of Approach (CPA) or time when fix is abeam the aircraft.	CPA is easiest to compute of the two methods, but can be misleading. If the task is to cross the fix before taking next action, such as a turn, the relative bearing of the fix from the aircraft is more reliable than CPA.
10. Holding Pattern		
Turn: (Outbound)	When aircraft is abeam the holding fix, START Turn Capture, Track, Rollout and Stop segments and logic.	The outbound leg desired heading should be adjusted for winds.
Outbound:	START tracking of outbound heading segment when Turn Stop logic is TRUE.	Note: These are "starter" algorithms; the pilot task is to adjust the outbound leg track across the ground for winds, or possibly a non-standard rate turn to the outbound leg. Also, this logic does not address hold-pattern entry. More complex logic may be required.
Turn: (Inbound)	When outbound heading track segment is TRUE, START Turn Capture, Track, and Rollout segments.	
Inbound:	Use Course Capture and Track logic.	
11. ILS or GCA Approach Centerline		
Capture:	When prior segment ends and Course error and error rate are converging in the proper direction, START.	Course error is assumed to be in degrees. Similar to VOR or TACAN course capture.

TABLE 3. EXAMPLE MEASURE SEGMENTATION LOGIC (Continued)

Segment	Start/Stop Logic	Comment
11. ILS or GCA Approach Centerline (Continued)		
Track:	When centerline error and error rate within limit cycle for 15 seconds, START 10 seconds earlier.	Nominal, trial limit cycle values are centerline error of 2 degrees and error rate of 2 degrees per minute.
Stop:	When Altitude is less than published Decision Height, STOP.	Note: One might wish to continue centerline measures through to touchdown, or missed approach measures.
12. ILS or GCA Glideslope		
Capture:	As dictated by prior segment and when Glide Slope error and error rate are converging in the proper direction.	Depending on procedures, the glide slope may be captured from above or below it, and this may be before or after the capture of GCA centerline or ILS localizer.
Track:	When Glide Slope error and error rate are within limit cycle for 5 seconds, START 3 seconds earlier.	Nominal, trial limit cycle values are Glide Slope error within 0.25 degrees and Glide Slope error rate within 0.025 degrees per second. Note: This will vary with aircraft stability and final approach airspeed, and will need fine tuning.
Decision:	When Altitude is less than Decision Height, STOP Glide Slope Tracking and measures.	
13. Landing		This logic assumes that the approach ends and the landing begins at the threshold of the runway, ramp of the aircraft carrier, or edge of the helicopter or VSTOL landing pad.
Start:	When passing over or abeam the approach end of the runway, START landing.	

TABLE 3. EXAMPLE MEASURE SEGMENTATION LOGIC (Continued)

Segment	Start/Stop Logic	Comment
13. Landing (Continued)		
Touchdown:	When weight on either main landing gear (or skid) for 1 second, measure touchdown and START Rollout.	If there is a bounce or skip, the first impact counts as the touchdown, but additional touchdowns should be noted.
Rollout:	START when touchdown TRUE; when Airspeed is less than 40 knots, STOP.	Note: There is no roll-out for helicopters unless a rolling landing is made; in that case the airspeed criterion will have to be adjusted. Neither does rollout apply to carrier landing.
14. Bolter		For carrier landing when the hook fails to catch a wire.
Start:	When tail hook passes the #4 wire, START bolter.	
Stop:	When weight off wheels, STOP.	Note: At this point it is assumed that acceleration and climb segments will start.
15. Missed Approach or Go-Around		
Start:	When aircraft "waved-off" by control authority, or go-around is initiated by the pilot, START.	Go-around procedures vary with aircraft. The typical sequence will be full thrust, arrestment of descent, retraction of gear, and acceleration to a climb airspeed and flap retraction schedule.
Stop:	When Altitude Rate is greater than zero, STOP.	Note: It is assumed that accelerate and climb segments will follow, per the desired procedures.

TABLE 3. EXAMPLE MEASURE SEGMENTATION LOGIC (Continued)

Segment	Start/Stop Logic	Comment
16. Formation Stationkeeping		
Capture:	When DX, DY, DZ and DX rate, DY rate and DZ rate converging then START.	DX, DY, and DZ are the differences in the X, Y and Z coordinate axes between the given aircraft and the target aircraft (formation leader). In some cases, a vector of all three (the slant range distance) is sufficient. The limit cycle will have to be determined empirically for each case; distance and rate of change will vary with the type of formation.
Track:	When DX, DY, DZ (and rates) within limit cycle for 15 seconds, START 10 seconds earlier.	
Depart:	When DX, DY, DZ (and rates) divergent outside of limit cycle for 15 seconds, START when divergence first occurred.	
17. Ground Attack Segments		For the roll-in to delivery, final dive, and recovery from entries using the box, cone and pop-up entry patterns. For more detail, see Vreuls and Sullivan (1982).
Roll-in:	When Absolute Value of Roll is greater than 80 degrees for 3 seconds, START 4 seconds earlier.	This is the roll-in to the final dive. The Roll-in segment may be subdivided for the pop-up maneuver. This logic will need to be changed to suit the delivery task; it represents a starting point only.
Roll-out:	When Roll-in is TRUE and the Absolute Value of Roll is less than 90 degrees, STOP Roll-in and START roll-out 1 second earlier.	
Tracking:	When Roll-out is TRUE and the Absolute Value of Roll is less than 15 degrees, for 2 seconds, STOP roll-out 2 seconds earlier and START final tracking.	
Release:	When trigger is pulled (or weapon is released), START and STOP release.	Release is a momentary event unless strafing is the task; if strafing, release STARTS as shown, but STOPS when trigger is released.

TABLE 3. EXAMPLE MEASURE SEGMENTATION LOGIC (Continued)

Segment	Start/Stop Logic	Comment
17. Ground Attack Segments (Continued)		
Recovery:	When g-force is greater than 2 gs for 5 seconds, START 6 seconds earlier. When positive climb rate is achieved, STOP recovery.	
18. Hover		
Capture:	When DX, DY, DX rate and DY rate are converging, START.	DX and DY are the positions in the X and Y coordinates between the aircraft and the hover point. Alternatively, the straight line distance (a vector of DX and DY) may be used.
Track:	When DX, DY and their rates are within the limit cycle for 15 seconds, START 10 seconds earlier.	The limit cycle values of DX, DY and their rates will have to be determined for each measurement situation.
Depart:	When DX, DY and their rates are outside the limit cycle and divergent for 15 seconds, START 15 seconds earlier.	
19. Translation To Forward Aerodynamic Flight		For helicopters and V/STOL aircraft.*
Start:	Use Hover Depart Start logic.	
Stop:	When Airspeed is greater than translational lift airspeed, STOP translation.	Translational lift airspeed will vary with specific helicopters, but it usually is about 20-30 knots. For V/STOL aircraft, use Nozzle Selector Lever position.

TABLE 3. EXAMPLE MEASURE SEGMENTATION LOGIC (Continued)

Segment		Start/Stop Logic	Comment
20. Translation From Forward Aerodynamic Flight			
	Start:	When Airspeed is less than translational lift airspeed + 5 knots, START translation.	For helicopters, deceleration logic should be used for the slowdown from cruise to the translational lift airspeed. For V/STOL aircraft, START when Nozzle Selector Lever is moved from full aft position.
	Stop:	Use Hover Track Start Logic.	

* For more detail on V/STOL see Hennessy, Sullivan and Cooles (1980), Ringland, Craig and Clement (1977), and Naval Air Systems Command (1975).

The power of using windows and a short time history is shown by the turn segmentation logic in Table 3. This logic assumes that a 15 second time history is maintained; flags are placed in the time history when roll angle and initial heading error are divergent (similar in sign) in the proper (expected) direction of turn. When roll attitude and heading error are greater than the limit cycle (10 degrees of roll and five degrees of heading) for five seconds, the logic looks back to find the first occurrence of continuous divergence, which would represent the time that the pilot first rolled into the bank angle to start the turn. We know of no other logic which can pinpoint the start of a turn this precisely.

Not all logic requires retroactive starting or stopping of segments. For example, the start of tracking takes place when certain conditions have been met for a period of time, as illustrated for turn, climb and descent, and acceleration and deceleration tracking. Also, some logic decisions can be made with simple tests, such as a reversal of bank angle (when changing a turn from left to right or the converse), or passing abeam a navigational fix.

There are some segmentation algorithms which are so dependent on the desired information (what one measures) that they do not fit the format of Table 3. For example, segments of aerobatic, basic fighter, and air combat maneuvers depend on the aircraft characteristics and what one wants to know. One way of segmenting and measuring these types of maneuvers is described in Appendix A.

COMBINING LOGICS

For a given flight profile, one would construct measure segment modules composed of combinations of the logics which are shown in Table 3. When combining logics, the measurement analyst may have to make decisions on which of two (or more) possible logics should control the measurement. For example, if measuring a climbing and descending turn pattern, one may have to decide whether the turn logic or the climb and descent logic is to control the maneuver segmentation. In typical patterns of this sort, a pilot is required to turn through a given number of degrees, and change from a climb to a descent in the middle of the turn; does the logic signal passing the mid-point of the turn, when the climb reverses to a descent, or both?

The answer to this question depends on the information which the user wants, and the interaction between the segmentation logic and measures which are taken. One can segment the turn task, looking for the rollout on the desired heading. Also, the climb and descent logic can look for the reversal from a climb to a descent, and measure the error from the desired heading at that point, as well as climb and descent performance independent of the turn performance.

If one constructs the segmentation logic and measures properly, it should be possible to compare similar task performance during various maneuvers which contain that task. For example, in climbing and descending turns, one can compare (a) climb and descent performance during a turn with climb and descent performance while holding a constant heading, or (b) turn

performance while holding a given altitude with turn performance while climbing and descending.

Such comparisons make it possible to diagnose performance to some extent. If a pilot's performance is poor during climbing and descending turns, but performance is good on constant heading climbs and descents, and constant altitude turns, it would be rational to conclude that the pilot just needed practice on controlling both axes at the same time (or for a non-training study, there was something in the independent variables that affected good control of both axes simultaneously). If, on the other hand, turn performance is poor in all cases, one could conclude that the pilot had not learned how to control a turn well, and one might want to provide additional training on controlling turns (or for non-training studies, there is something in the experiment that affects turn performance).

These kind of comparisons are possible only if the segmentation logic and measures are exactly the same under all conditions. Similarly, a comparison of the data from one study to the next would be aided if the measures were taken with the same segmentation logic. Seldom, however, does this occur; typically, measures are controlled by the experimenter or project pilot, or are constructed from "cleaned-up" time history tapes after the data are collected for practical reasons.

No one, however, has suggested using a standard method to control measure segmentation. It is doubtful that measures can be standardized for all research purposes, but perhaps the use of standard measure segmentation logic can emerge from attention to these issues, test and refinement of the suggested segmentation methods.

SECTION 6.

TRANSFORMATION GUIDELINES

Measurement should provide needed information. At the outset, then, the question should be asked: "What is it that we wish to know as a result of measurement?" A systematic approach can be developed from this, but since the number of such questions is beyond our ability to treat here, measurement will be treated in more general terms.

Our emphasis will be on general characteristics of data which can be described through the use of mathematical transforms. Specific computer programs also will be presented. It is believed that the reader will be able to translate information needs into the terms presented here, and thereby be led directly to the tools for answering application-specific questions.

The level of discussion in this chapter will be in terms of a range of potential data transforms. The concept of mapping is fundamental to this approach. A mapping is a rule for placing the members of one set into 1:1 correspondence with the members of another set. For example, the average (or mean) is a rule for mapping from a set of numbers into a single number. This is the case of many elements being mapped into one. Such mapping is the basic purpose of the measurement process and therefore the topic of this chapter.

Although it is thought that the ideal of measurement is to reduce the complex down to a single number, one should be reminded of the purpose of measurement to provide information, and go on to consider the limited information conveyed by a single number. Rather than dwell on mapping of "many to one," therefore, and in the interests of general utility, the mappings ordinarily will be from one set to another full set, or "many to many." The result of transforming will be a large set, such as a curve; never-the-less, specific variables of the transformed set will be identified, allowing simplification. By picking specific variables of the transformed set, the end result can be a mapping of many to one (or few).

What, then, is gained by many-to-many mapping? Assuming that we have been recording some continuous variable, we have what may be displayed as a wiggly curve plotted against a time base. Examination of the wiggly curve may not be informative in itself. After transforming, however, we can examine the data in a new "domain," where the curves have new interpretation.

For example, as a result of transformation, the data are now a distribution of amplitude, or the frequencies contained in the original data, the interpretation may provide new insight into the nature of the "wiggly lines". We might say that we have transformed into a new domain with dimensions that impart insight and meaning. There are no guarantees that transforming will be successful, and certainly no magic is involved; but the process has been worthwhile in the past, and is the foundation for much current measurement.

The remainder of this section will discuss a number of transforms which have had past utility or show promise for the future. We will begin with a discussion of measures for time history data (i.e., "wiggly lines" inked on a stripchart recorder). The subsequent discussion will include amplitude distributions, Fourier transforms, simple human operator models, and reference frameworks for interpretation. Associated with each of these areas of emphasis will be summaries which should be useful for reference purposes. A collection of FORTRAN programs is included in the Appendix B.

GENERATION OF SAMPLE DATA

In order to provide concrete examples and a means for demonstrating useful computer programs, some example data were generated. Rather than collect data from real human performance, a computer program THDAT was written to provide test data in a form for input to other measurement and analysis programs. The advantage is that the data characteristics are known, so measurement results are interpreted more easily than with data of unknown characteristics.

A simple control system is assumed. The human operator has direct control over a pointer or aiming device, and the task is to keep the pointer aligned with a moving target; thus, the human is to operate a simple position control system. Consequently, there are three signals to monitor: the target motion (TGT), the human operator response (HR), and the error (ERR) which is the difference between the other two signals.

The target is given an oscillatory motion, which is generated by the sum of a number of sinusoidal components (0.1, 0.2, 0.3, 0.4, and 0.5 Hertz). To make the motion something which the human operator can track, the frequencies are held within manageable limits (much less than 1 cycle (Hertz) per second) and the amplitudes are reduced with frequency. Also, in keeping with some of the characteristics of the human operator, two types of delays are included: (a) a time delay, with the human operator response signal operating about 0.2 second behind TGT, and (b) a phase delay, with lagging phase angles introduced into the higher frequency components. Additionally, a small noise component is added at a high frequency which could be due to tremor or dither in the human operator control response. In all, THDAT produces one minute of data for one trial.

To show the final result, the data were processed by a computer program LSTRP3 which plots the data in strip chart format. LSTRP3 is a utility program which plots on a printer capable of printing in excess of 100 columns across the width of the page. Since this capability is not available on some printers, and since the strip chart plot is a useful utility, a version which plots just one signal on 80-column paper (STRIP) also is presented in Appendix B for reference.

TIME-HISTORY MEASURES

Information Produced. Specific characteristics of continuous time history recordings are mapped into specific measures such as peak value or average value.

Definition. Strictly speaking, there are no mathematical transformations corresponding to this class of measures, but we call these measures "transforms" for purposes of describing measurement structure. We start out simply by picking out interesting characteristics from the wiggly lines (cf., output from LSTRP3 or STRIP). The following is a narration of a list of possible measures which may be used to describe such wiggly lines.

Specific measures: Program THMEAS, in Appendix B, was developed to illustrate time-history measure programming and to provide ready-to-use tools.

Time on Target (TOT). The test data were generated under the assumption of a task involving pointing at a moving object, such as aiming a gun at a moving target. Some of the early World War II tracking research was performed for such gunnery systems, and a common measure was the time the operator tracked on target. A tolerance band is established and a clock started whenever the aiming point was within this band, and shut off whenever outside. Given digital data, the measurement of TOT can be performed with a single FORTRAN IF statement.

Zero Crossings (ZCROSS). TOT provides a measure of accuracy, but one also may want some measure of stability, or the smoothness of tracking. A simple measure of smoothness can be obtained by counting the number of times (per unit time) the aiming point crosses the center of the target. For example, the human operator may obtain a large TOT but if the ZCROSS measure also is high, it will be apparent that the aiming point was sweeping rapidly across the target.

Peak Value (PEAK). It also may be of interest to determine the peak value of error which occurs during tracking to provide boundaries on the performance. The peak value is computed by examining each data sample and storing it as PEAK if it is larger than the value previously stored as PEAK. An extension of this measure could be provided by computing both positive maximum and negative maximum values. The example program only stores the greatest absolute magnitude without concern for sign.

Average Error (AE), Absolute Average Error (AAE), Mean Square Error (MS), and Root Mean Square Error (RMS). In addition to the boundaries of performance, one may be interested in average performance in the form of average error. There are a number of average error measures appearing in the literature. The most simple and straight-forward form is AE which is the sum of error (for each sample of data: error = target position - human operator response) divided by the number of samples. It will be seen that error is a signed quantity, and in summation for computing the average positive and negative errors cancel. Consequently, poor performance can result in zero value for AE.

One correction is to take the absolute value of error before summation, and this results in the AAE score. Another approach with some precedence in engineering and statistics is to square error before summation, and this results in the MS score. A further variation is to take the square root of the MS score, resulting in the RMS score.

While the difference between the AE score and the others is clear, the difference between AAE, MS, and RMS may not be clear. The basic consideration is the effect of squaring error before summation, because squaring affects large values and small values disproportionately; that is, with AE and AAE the effect on the end result is in proportion to the magnitude of error, but with MS and RMS large values are weighted more heavily.

For example, with error values 1 and 2 the sum is 3, but the sum of the squares is 1 + 4 = 5. RMS involves a final square root operation, but the disproportionate summation is still present. The ratio of RMS to AAE, for example, is not fixed, but depends on the shape of the signal. In a given study, if the wave shapes stay fundamentally the same, there will be a fixed ratio and there is no point in measuring both AAE and RMS; however, if the behavior and the waveshape changes, then the ratio of RMS to AAE will vary, and different results could be obtained depending on the measure selected.

Reversal Count (REVERS). A basic principle in tracking tasks is that the operator should always act to reduce error. Any behavior which operates to increase error is noteworthy (some aircraft instruments, for example, may promote error increase). Whenever the operator overtly acts to increase error, such an action may be called a control reversal.

Reversal measurement involves instant-by-instant evaluation, comparing the direction of control movement with the direction of error. A simple algorithm for measurement of reversals is to examine error to determine if it is increasing in magnitude (positively or negatively), and to determine if the operator control action is moving in the direction of the error increase. The FORTRAN code for REVERS will be found in THMEAS in Appendix B. A computer run of THMEAS yields the scores shown in Table 4.

TABLE 4. TIME HISTORY MEASURES

Measure	Value
Time (Samples) on Target	105.00
Zero Crossings	54.00
Average Error	0.02
Absolute Average Error	2.28
Mean Square Error	9.77
Root Mean Square Error	3.13
Peak Value	-9.29
No. of Reversals	84.00

AMPLITUDE DISTRIBUTION TRANSFORM

Information Produced. The amplitude distribution transform divides the signal amplitude range into a number of bands, and counts the occurrences within each band. One obtains the frequency of, or the percent of time within, each band during the recording of a specific continuous variable. Note that the sequential or temporal characteristics of the original signal are lost through this transformation; for example, a waveform with multiple peaks and valleys could have the same amplitude distribution as one with one peak and valley. The result is a plot, and from this graphic presentation, one may determine a number of specific measures to be appropriate in describing the plots. The amplitude distribution, per se, may therefore be considered a preliminary data scanning tool.

Definition of the Transform. The expected range of the continuous variable is divided into equal-sized bands. The variable is sampled at equal time intervals, and a band counter is incremented for a band as each time-sample is tested. The resemblance to the TOT score may be apparent, and the transformation may be viewed as scoring of multiple TOT tolerance bands.

There are two forms: the amplitude distribution and the cumulative amplitude distribution. The procedure just described will produce an amplitude distribution. Instead of incrementing just one band counter for each data sample, if one increments all bands representing values equal to or greater than each sample value (doing this for each data sample in turn), then the cumulative form of the distribution results.

The amplitude distribution corresponding to the test data is presented in Figure 3 and the cumulative distribution for the same data is presented in Figure 4. The corresponding FORTRAN programs, AMPDIST and CUMDIST, are available for reference in Appendix B.

From either of these presentations, one can readily identify the smallest value, the value for which 100 percent of the data was smaller (or, in other words, the largest value), and the value for which half of the data were smaller and half larger (the median value). The median value is more precisely read from the cumulative distribution. The cumulative distribution has the distinct advantage of being less sensitive to the size of amplitude bands selected. If the amplitude bands which are selected are too small, then a rather ragged plot results. The selection of amplitude band size often is a tedious iteration.

Specific Measures. The amplitude or cumulative distribution ordinarily would be produced to see the shape and general characteristics of the distribution. The distribution could be bi-modal, or highly skewed. Perhaps after such preliminary visual screening, other computer programs could be used to determine the specific characteristics of the distribution.

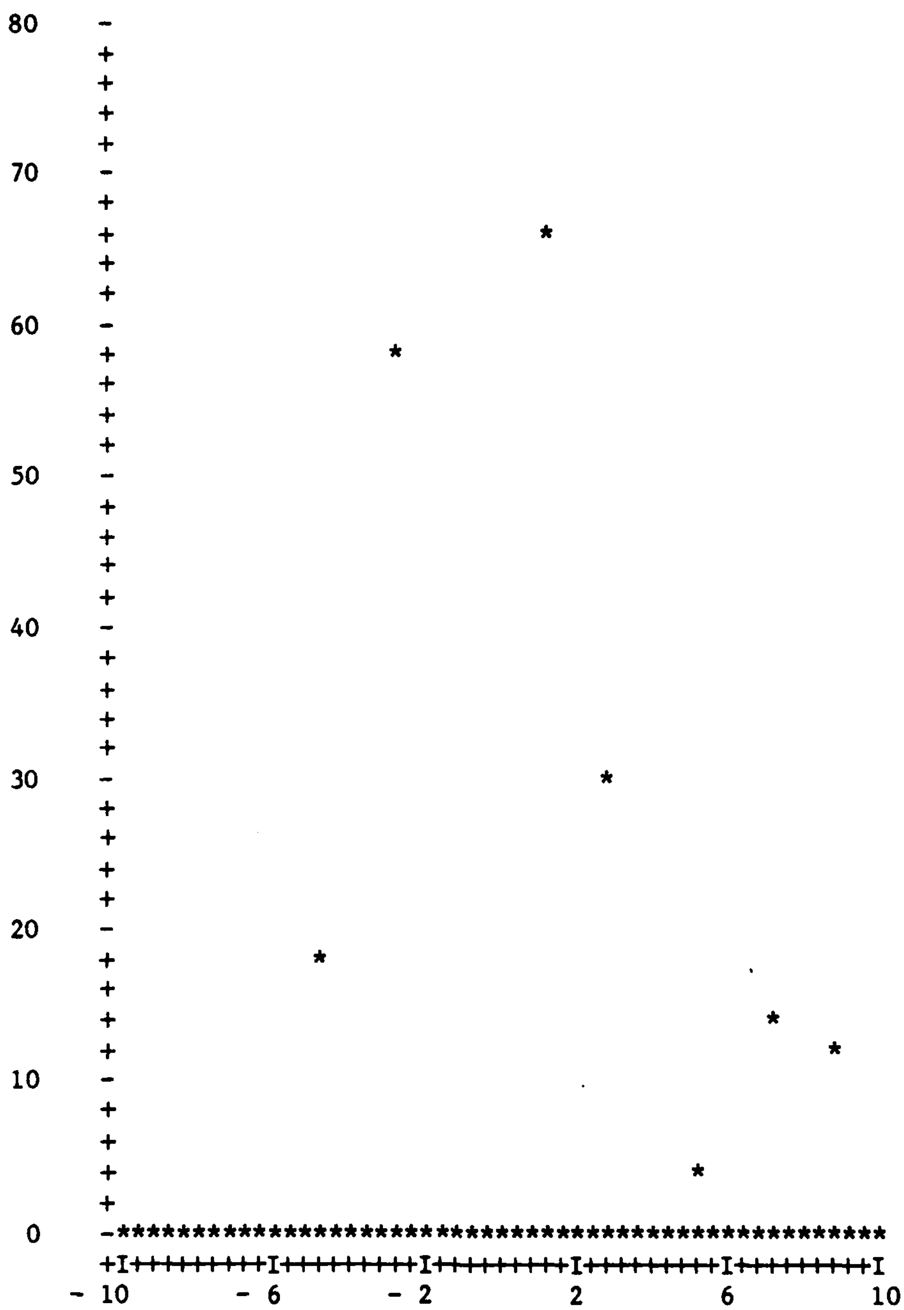


Figure 3. Amplitude Distribution for Example.

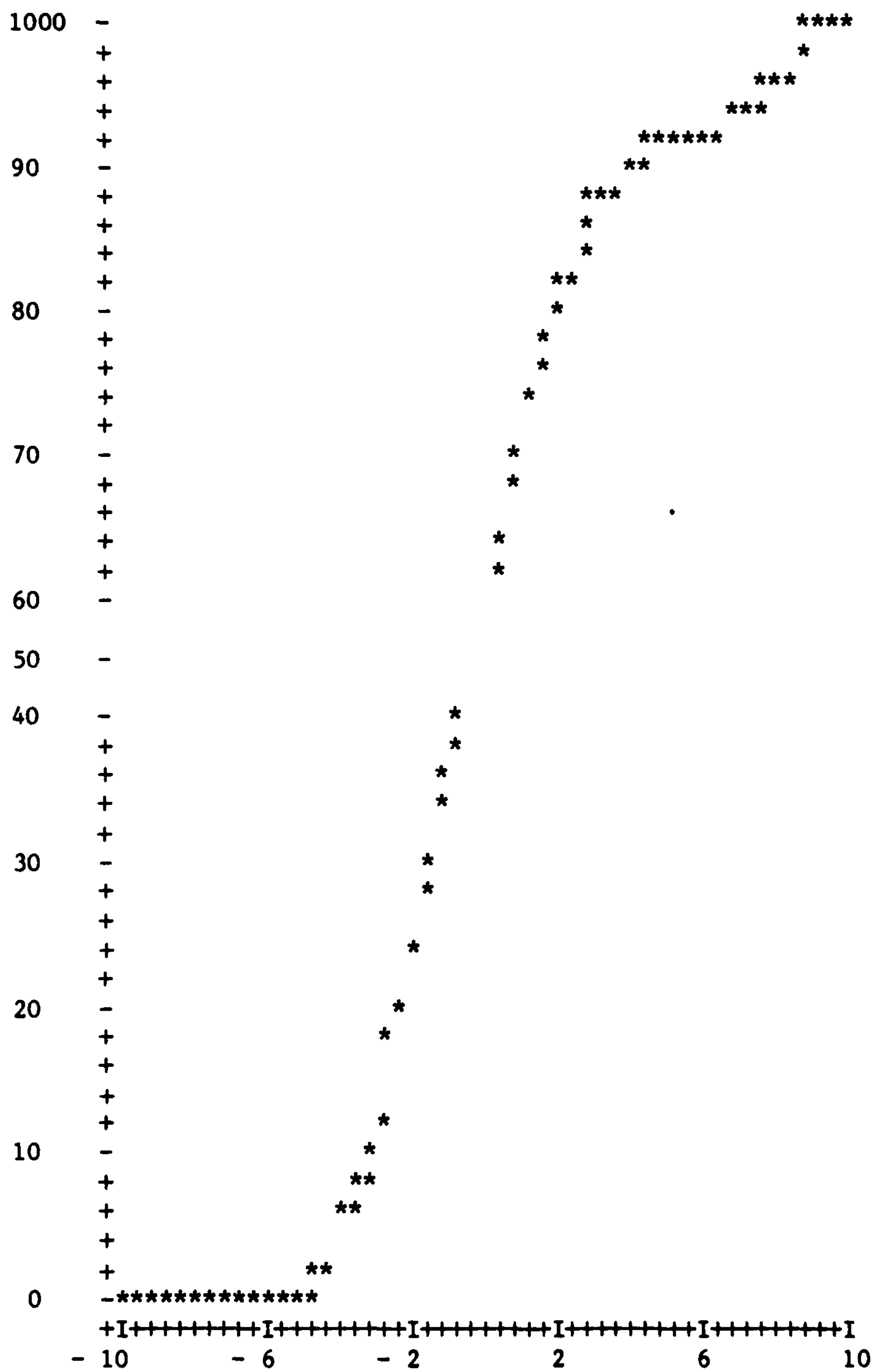


Figure 4. Cumulative Distribution for the Example.

For example, program MISR is provided in Appendix B to measure the moments of the distribution. If the distribution is normal, measurement of the mean and standard deviation are in order; note, however, that the mean was discussed with time-history measures, and that the RMS provides a measure of the standard deviation. If the distribution is not normal (many human performance distributions are not normally distributed), then the median and quartile range may be the measures of choice. The distribution of absolute error for a tracking, or error nulling task, for example, often follows a Poisson distribution, in which the mean and standard deviation are nearly the same; in this case, measures of both would be redundant.

It is important for measure selection to know the characteristics of the amplitude distribution of variables of interest; this step in measure development is overlooked frequently, but should not be.

FOURIER TRANSFORM

Information Produced. The Fourier transform provides a representation of the original signal in terms of sinusoidal frequency components, each component being described in terms of an amplitude and phase angle characteristic. This representation for human control performance provides a basis for discussion in control engineering terms, in the same language as is used for describing and analyzing the rest of the system.

Definition of the Transform. If the given signal is periodic, that is, it repeats every T seconds, the Fourier theorem states that the signal can be represented by a (possibly infinite) number of sinusoidal components (harmonics) with frequencies 1/T, 2/T, ..., n/T Hertz (cycles per second). That is, the signal can be represented in a form similar to that used in program THDAT to generate example data.

The formula is: f(t) = a0 + summation[an*cos(wnt)+bn*sin(wnt)] where w=2*pi*(1/T) and n=1,2,....

If the given signal is aperiodic, a similar representation is possible, but the component frequencies are infinitesimally spaced, resulting in a continuous spectrum. The resulting function is referred to as an amplitude density spectrum, and a related function is the spectral power density. Note that for periodic functions one obtains what is called a line spectrum; that is, rather than a continuous spectrum, a series of discrete harmonics plots as a equi-spaced series of vertical lines.

Specific Measures. A computer program suitable for performing a Fourier transform of the example data (PFOURIER) is included in Appendix B, and the results are tabulated in Table 5 and plotted in Figure 5.

TABLE 5. FOURIER COEFFICIENT

FREQ	COS	SIN	AMPL	PHASE
.000	-.025	.000	.025	90.
.016	-.051	.020	.055	-68.
.033	-.051	.031	.060	-59.
.050	-.052	.054	.076	-44.
.066	-.055	.097	.112	-30.
.083	-.061	.202	.211	-17.
.100	-.665	9.789	9.812	-4.
.117	-.039	-.175	.179	-167.
.133	-.047	-.059	.076	-141.
.150	-.053	-.006	.053	-97.
.167	-.058	.046	.075	-52.
.183	-.072	.155	.171	-25.
.200	-.660	4.827	4.872	-8.
.217	-.027	-.213	.215	-173.
.233	-.044	-.094	.105	-155.
.250	-.054	-.041	.068	-127.
.267	-.064	.008	.064	-83.
.283	-.083	.104	.134	-39.
.300	-.612	2.878	2.942	-12.
.317	-.027	-.232	.233	-173.
.333	-.055	-.123	.135	-156.
.350	-.075	-.079	.110	-137.
.367	-.103	-.045	.113	-114.
.383	-.174	.010	.174	-86.
.400	-1.872	1.109	2.176	-59.
.417	.115	-.164	.200	145.
.433	.022	-.103	.105	168.
.450	-.016	-.078	.080	-169.
.467	-.051	-.060	.079	-139.
.483	-.123	-.035	.127	-106.
.500	-1.422	.314	1.456	-78.
.517	.160	-.102	.190	122.
.533	.071	-.075	.104	137.
.550	.042	-.065	.078	147.
.567	.028	-.059	.065	155.
.583	.018	-.054	.057	161.
.600	.011	-.050	.052	168.
.617	.005	-.047	.048	173.
.633	.001	-.045	.045	179.
.650	-.003	-.042	.042	-175.
.667	-.007	-.040	.041	-169.
.683	-.013	-.038	.040	-161.
.700	-.020	-.036	.041	-151.
.717	-.032	-.033	.046	-136.
.733	-.063	-.028	.069	-115.
.750	-.480	.015	.480	-88.
.767	.084	-.042	.094	117.
.783	.036	-.036	.051	225.
.800	.022	-.034	.040	147.
.817	.014	-.032	.035	155.

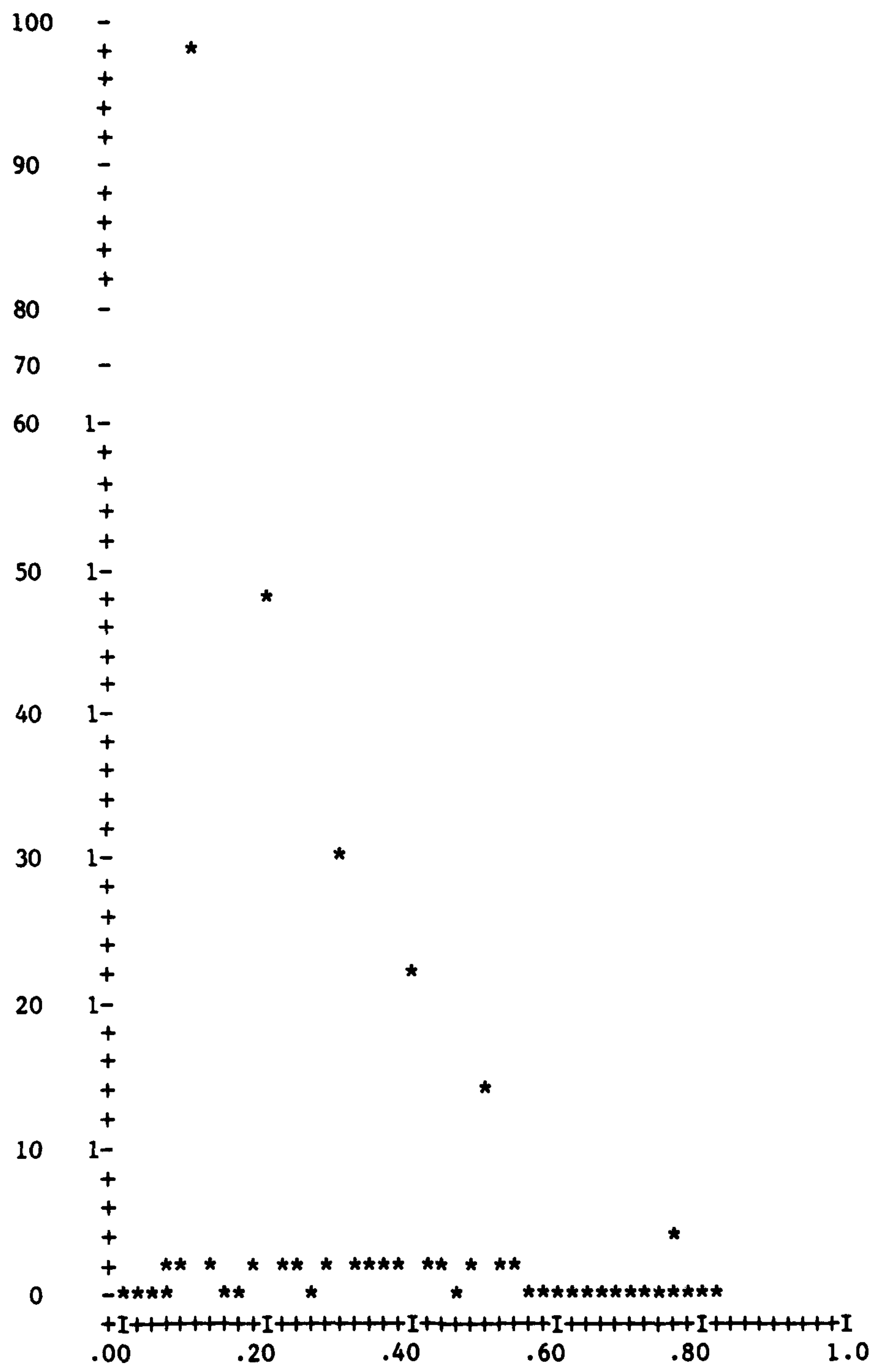


Figure 5. Fourier Transform Plot for the Example.

Transfer Function. If a Fourier transform is produced for the input to the human operator (for the example, this is a Fourier transform of the target motion), and if a Fourier transform is produced for the human operator response, the input/output relationship of the human operator can be examined on a frequency-by-frequency basis. If this is done for the example, it will be seen that there is some attenuation and phase delay for the higher frequencies, and the introduction of a high frequency not present in the input (noise). This approach is the basis for many control theory models of the human operator, including the quasi-linear transfer function model (McRuer and Krendel, 1957).

Band Width. Consideration may be limited to just the frequency band in which spectral power is concentrated. This is measured as the frequency band between points on the spectrum at which power is down one half the central or nominal value (or what is equivalent, the points at which amplitude is 0.707 the nominal value). Such data are often useful in the engineering of the control system.

Signal and Noise. As pointed out in the example, additional power (or noise) can be determined with the Fourier transform; the amount of noise, or the ratio of signal to noise is also of value in system assessment and design.

Data Collection Considerations. The sampling a continuous signal, and then using the sampled data for analysis, can introduce artifacts into the frequency analysis. A problem, called aliasing, is introduced when sampling is performed at a rate less than twice the highest frequency component in the signal (see Blackman and Tukey, 1958). That is, if the highest frequency component is 1 cycle per second, such data should be sampled at least twice per second. This criteria, while developed for Fourier transforms, is often used in practice for the sampling of data for other uses.

INPUT-OUTPUT RELATIONSHIPS

The previous sections generally have addressed the measurement of one variable at a time. Where the human operator's response is the result of specific stimuli, however, the relationship between display-input and operator-output can be of more interest than either alone. This subsection will examine some basic tools for analysis of relationships in the context of models of human operator input/output (e.g., human operator transfer functions). This is an extensive subject which will not be pursued comprehensively; in the following, a rather simplified approach will be taken just to provide basic illustrations.

Linear Regression. A possible model for human operator control is that the human output is proportional to the displayed input (to the extent that human output can produce such a response accurately). One might have to augment a linear response with time delay and lags at high frequencies to produce an output similar to that which is actually produced by human operators. For the example data which we have been using we know that there are delays and lags, but otherwise a reasonable approximation to a proportional response is exhibited. In fact we are aware that the data

were created with a time delay, roll-off at high frequency, phase lag along with the roll-off, and high-frequency noise.

The specific procedure is to use a regression analysis to find the equation y = b0 + b1 * x where y is the human operator's response (HR) and x is the stimulus signal (ERR). The delay factor is incorporated by shifting the data so that previous stimuli are paired with latter responses. In brief, the procedure is:

1. Shift data so that input sample n corresponds to output sample n+1, input sample n+1 corresponds to output sample n+2, etc.

2. Perform a regression analysis and determine Rsquared (the amount of variance accounted for).

3. Repeat steps (1) and (2) until sufficient analysis has been conducted to determine the shift (time delay) which yields the maximum Rsquared.

The model being suggested at this time is simply a proportional response with a time delay. The proportionality constant (or gain) can be determined with a regression program like MULTR, and the time delay found by shifting the data samples for human response so that samples are being correlated with input target samples one, two, or three samples earlier in time (up to a few tenths of a second shift) to find the shift which provides maximum correlation. A program like CROSS can also be used for this purpose. If there are multiple inputs to which the human operator is responding (e.g. altitude, rate of climb, acceleration, etc.) then a multiple regression program like MULTR should be used to find the best linear combination of inputs to predict the operator's output.

Describing Function Model. Another approach, previously alluded to, is to examine the time delay, gain and phase shift for each input frequency (five components in the example data) and to combine all other aspects of the operator's response into a noise, or remnant, term. This is the approach taken in many well-used and successful control engineering models (cf., McRuer and Krendel, 1957). In brief, the approach is to:

1. Determine the magnitudes and phase angle of specific frequencies in the input signal.

2. Determine the magnitudes and phase of the same frequencies in the operator's output.

3. Compute the ratio of the magnitudes and the difference in phase angles.

4. Fit a polynomial equation to these data.

The purpose of this discussion is to suggest possible transformations and measures which may be applicable to some unknown specific application. For example, for a particular case, one may be interested in only the point

at which roll-off occurs, or the nominal gain, or the time delay, or just the noise component. Note that a Fourier, or sinusoidal, representation must be possible for the input signal; however, this situation can be created in signals generated in the laboratory, and for many non-stochastic signals in natural environments.

Optimal Control Theory. Various performance criteria can be applied to a given control task. For example, it may be desired to achieve a specific goal in the minimum amount of time; it may be desired to achieve the objective having used the minimum amount of fuel; it may be desired to minimize the magnitude of certain system states.

An optimal control policy is one which minimizes some performance index within whatever constraints may be placed on the control exerted (e.g., control is finite, limited by available control movements and available power). It may be seen, therefore, that there is no universally optimum control policy; it depends on the specific performance index which is defined.

Optimal control theory is of interest for describing human performance for at least two reasons: First, the optimal control for a specified performance index can serve as a standard of comparison for the control behavior exhibited by the human operator; this point of view will be discussed in the next section on interpretive frameworks.

Second, optimal control theory can be used to determine the performance index which is optimized by the exhibited human control behavior. Any control behavior can be viewed as optimum for some performance index. If that performance index can be identified, one is provided with a mapping of control behavior into a set of performance index parameters, such as weightings of control and vehicle states for which the control exhibited is optimum. The use of performance data and optimal control theory to find the performance index being optimized has been called the "inverse optimal control process."

This approach was identified by Obermayer and Muckler (1965) for the selected class of linear systems and quadratic performance indices. The approach was used effectively by Connelly (1977), wherein performance was pre-computed for various performance indices, and then specific control behavior was matched with this set to determine the approximate performance index for which the behavior was optimum.

The determination of the inverse optimal control policy can be a major challenge except for specific classes of control problems, however, the Connelly method can be applied to ordinary experimental data. Briefly, the procedure is as follows:

1. Calculate the performance which would result from a broad selection of performance indices.

2. Compare the performance resulting from each of the results of (1) to the specific experimental conditions of interest (e.g., each subject experimental trial, each display design, and so forth).

3. Pick the performance index which yields the best match.

4. Test the representation using a regression analysis performance of the selected model to predict performance across subjects.

INTERPRETIVE FRAMEWORKS

State Space Representation. A physical system can be described in terms of a number of variables, so that specification of these variables at any time fully define the system state so completely that, with the appropriate model, future states could be predicted. If each state variable is then taken as the axis of a coordinate system, the complete history of a system can be described as a path in this state space.

The state space can have a large number of dimensions, but a useful analysis tool is the phase plane, which plots position (x) against velocity (xdot), with x being a selected state variable.

State space and phase plane representations can be useful transforms for manned systems. For example, for aircraft landing, one may plot altitude (position) against rate of descent (velocity) using program PLOT to provide an informative look at the landing profile. Such a look may suggest specific measures of interest, such as the rate of descent at several specific altitudes.

State Space Cells. While each dimension in state space may be a continuous variable, it may be useful to divide each dimension into parts (e.g. high-, medium- and low-altitude), and thereby divide the space into a number of cells. If the division is done properly, each cell will contain all performance which can be considered the same and is treated in the same manner by the analyst. Several examples should serve to illustrate this technique.

1. Phase Plane Tolerance Measure. Consider a plot of error on one dimension of a phase plane, and the derivative of error (error rate) on the other dimension. One may draw a tolerance circle, centered at the origin, which divides the space into two regions: outside the tolerance circle is the region of unacceptable error and error rate; inside the tolerance circle is a region of acceptable performance.

It may be noted that when error and error rate are opposite in sign, the operator is correcting the error (reducing the amount of error), and therefore is exhibiting desirable behavior. Alternatively, one may wish to count performance as being poor when it is outside the tolerance circle and also in one of the quadrants where error and error rate are the same sign (there is error, and it is increasing).

2. Air Combat Maneuvering Space. Air combat tactics often are discussed in terms of range, line of sight (angle off other aircraft off the nose axis) and aspect angle (angle off the tail axis of the other aircraft). Consequently, a three-dimensional space can be created using each of these variables as an axis. Further, regions of range, specific

regions of line of sight and aspect angle define conditions where it is appropriate to launch weapons, take evasive action, and execute specific offensive maneuvers. Using these variables as a guide, it is possible to divide the space into cells which represent specific types of tactical challenges.

3. Surface Ship Maneuvering. Consider the task of maneuvering a ship through heavy ship traffic without colliding with other ships. The primary variables of interest are (1) the closest point of approach (CPA) to each ship when the paths of each ship are projected into the future, and (2) the time to the closest point of approach. These variables can be further divided to regions (0-1.5, 1.5-2.0, 2.0-2.5, and above 2.5 miles for the CPA; 0-5, 5-10, 10-15, 15-20, 20-30, and above 30 minutes for the time to CPA). Consequently, the state space (plane) is divided into cells which can be used for analysis. For example, performance can be viewed as the process of transitioning from one cell to another, and one can measure the probabilities of these transitions.

4. Low Level Flight. Consider the task of flying a low level or NOE flight; the centerline of the desired course is plotted on a map, but it may not be necessary to be on the course centerline at all times, especially if the pilot knows his current position and is flying directly toward the next checkpoint; typical low level flight performance weights fuel consumption and time crossing each checkpoint heavily. A state space similar to the one described above for ship maneuvering can be constructed for the projected CPA at the checkpoint versus miles to go, and transitions to (and occupancy within) acceptable states may provide more satisfactory descriptions of pilot strategy and performance than centerline error.

5. Transitions. Frequently, there are families of acceptable curves of performance when transitioning from one steady-state to the next. For example, when turning over a checkpoint to capture a new outbound course there are various school solutions, which dictate the angle of bank to use and the desired intercept angle. These rules vary with the situation, but one can measure error from the ideal, or the parameters of a control law model of the transition.

Alternatively, one may simply score performance in regions of the phase plane using the rationale that as long as the pilots are converging on the course within acceptable boundaries of error and error rate, they are performing properly. If the convergence is too rapid for the situation, the aircraft is likely to pass through the course centerline and limit cycle, and be "caught" for being in an area of divergence. If pilots are not capturing the course quickly enough, a "maximum" time allotted for the transition can capture that fact. These methods can be used as well to score convergence on (or divergence from) ideal parameters for weapons release during air-to-ground and air-to-air weapons delivery tasks.

Performance Boundaries. In some cases the requirement for measurement computation is to derive complex criteria which is then used to assess raw untransformed performance variables. A case in point is the use of energy maneuverability diagrams for air combat maneuvering (Pruit and Maroney,

1980). The ability to out-turn an opponent is of paramount importance in air combat maneuvering; therefore, these diagrams show turning rate as a function of velocity (airspeed or Mach number). While performance may be assessed on this plane, the diagrams are drawn only for a specific altitude, requiring a three-dimensional space, but pilots typically view this as a plane with changing characteristics. A sample of this diagram (Figure 6) is shown below:

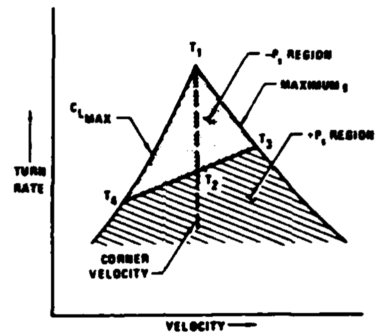


Figure 6. Key Turning Conditions on Turn Rate-Velocity Diagram.

Performance is bounded on an energy maneuverability diagram by two curves: one which reflects the maximum possible lift, and the other which reflects the maximum load (g's) which is allowed by the strength of the airframe structure. The point where these curves meet (T1) indicates the maximum rate of turning (called turning on the "corner"); however, since this point also represents a region of high drag, the maximum turn rate cannot be sustained.

The two bounding curves can be overlaid with a third curve (T4, T2, T3) which represents performance in which thrust (at military power) and drag are balanced. Below this curve there is an excess of power available and above it deceleration and/or loss of altitude must occur. T3 denotes the maximum turn rate which can be sustained; T4 denotes the minimum turn radius which is possible; T2 is the maximum sustained turn rate at the corner velocity (allowing momentary excursions to the maximum turn rate). Another consideration is that the curves of the energy maneuverability diagram must be re-computed for each change in altitude.

Performance can be measured as the error between actual conditions and optimal values, and the time required to achieve optimal. Furthermore,

where two dissimilar aircraft are engaged in battle, two sets of curves can be overlaid and tactics can be assessed by noting whether the flight variables favor one aircraft or the other.

The point here is that performance variables (velocity, turn rate, g's, altitude) are not transformed; it is the criterion which presents a problem for measurement processing. Once performance is presented within the energy maneuverability diagram, the raw data are imparted new meaning and additional transformations can be generated (e.g., measures of nearness to optimal conditions, and time to attain changed conditions).

Cost-to-Go. It is commonplace to establish a fixed path in space as a reference path in space and to measure all deviations from this path as error. For example, the glideslope and localizer beams establish a fixed path for instrument landing; it is easy to measure glideslope and localizer deviations, and to use these as a basis for measurement computations.

It does not necessarily follow, however, that once displaced from an nominal path in space, the best performance strategy is to return immediately to the nominal path. Once displaced from the path, another optimal path may be established, and further, if the operator is making correction, one may not wish to excise a performance penalty for the deviation which exists at the time.

Following techniques developed by Connelly and Zeskind (1975), it is possible to establish an optimal path for all points in space, not just a fixed nominal path. Given a performance index, its value (which can be viewed as a cost function) can be established as the minimal cost to go to the desired goal. This may be subtracted from the cost-to-go calculated for the operator's performance yielding only the excessive cost (the non-optimal part). Not only does this permit comparison to a reference which always reflects optimal performance, it also provides an instantaneous performance measure, permitting one to easily determine where critical instances of poor performance occurred. The instantaneous cost-to-go measure, therefore, can provide a basis for diagnostic measurement of the operator's performance.

GUIDELINES FOR MEASUREMENT SELECTION

Unfortunately, there are few concrete solutions to human performance measurement problems. On the other hand, one should not conclude that measurement selection and design is totally arbitrary. The problem is that information needs are specific to each situation, and the kit of tools is extensive, creating a measurement design problem with many dimensions.

The taxonomy of human performance measurement has not been detailed. Consequently, it is not possible to offer a comprehensive and structured discussion of measurement design. It is possible, however, to discuss a number of measurement applications which collectively involve a range of measurement tools which can be adapted to a majority of measurement problems. This is the approach to be used in the following paragraphs.

An application will be defined, the desired information identified, and a possible avenue to measurement will be discussed (one which is believed to be well suited to the application). Thus, for any application, the measurement designer may view the problem as a combination of several of the examples which are presented here and construct an appropriate composite set of measures.

Data Exploration. Perhaps one of the most difficult applications, insofar as prescribing a specific measurement approach, is the case where one is exploring experimental data in the attempt to discover unknown characteristics and relationships. One may be uncertain about the quality of the data; therefore, it is desirable to process data to develop some confidence in the usability and reliability of the data.

Data exploration measurement commonly includes time history measurement and plots of amplitude distribution for individual subjects and trials, as well as averages across trials and subjects. Outlying data points, multimodal distributions, and data variability should be evident from this data treatment. Confidence about data quality may be gained if some measures are collected where some characteristics of the data are predictable, and if, of course, the expectations are confirmed.

Since data characteristics and relationships often are perceived more readily in pictorial form than lists of numbers, the use of automated plotting routines can be beneficial. If the data characteristics are unknown, it is difficult to be more specific than this; otherwise, if the problem matches prior investigations to any degree, then the work of other investigators may help point the way.

Measurement to Derive Information for a Specific Class of Users. If the purpose of measurement is to derive information for specific people, for example presentation of student performance to a military instructor, then the information is largely defined by that which the instructor and student needs and will understand. Measurement design should be preceded by an information requirements analysis.

It should be clear that the primary displayed information must be in terms that the user can understand and apply; ordinarily, this is the measurement of time within tolerances, snapshots of state variables at critical instances, and identification of common student errors. One may wish to consider esoteric forms of performance measurement for the support of automated training functions as well as for information for display to the user personnel; however, in the latter case the training implications may have to be translated into statements in the language of the user.

Fixed Profiles. The measurement of performance for vehicle control along pre-specified profiles (paths) in space is perhaps the situation with the most extensive history. Amplitude distribution measures and snapshots of profile variables has been the most extensively used measures. One should note that these do not yield much information about human control behavior, and that human operator models, optimization theory (e.g. performance index optimized), and further segmentation (e.g., separate analysis of path capture and tracking) should be considered.

Tactical Maneuvering. The maneuvering of a vehicle in combat, where the consequences of actions are uncertain and strategies vary from moment to moment, is an area which is not fully understood by the experts, and what is understood may not have been well articulated. The measurement designer is understandably on unsure ground; however, two suggestions have been made earlier in this section. One is the development of complex criteria such as energy maneuverability diagrams; the other is the use of state space measures and the subdivision of state space into cells where measurement is more tractable (i.e., divide and conquer).

Control System Design and Analysis. There are well established methods for the design and analysis of automated control systems, and these have been adapted for use with manual and semi-automated control systems. It follows, to be useable, that measurement must conform with the existing design and analysis methods. Consequently, for these applications, some form of human operator control model must be employed. Some simple models were mentioned earlier in this section. Fortunately, simple and approximate models of the human operator and the machine subsystems often have been successful for control system analysis.

Display Evaluation. Measurement for the evaluation of alternative display designs is facilitated by information about the expected differences in performance to be produced by the displays. The expected differences in performance can include operator control behavior, operator decision behavior, and mission performance. Each of these represents a different measurement problem which has been previously discussed. Variations in operator control behavior should be reflected in the parameters of an appropriate model. Decision behavior may be reflected by transitions in state space or by changes in the performance index which is optimized. Mission performance can be expressed in some form of amplitude distribution measure.

Performance Diagnosis. Diagnosis, such as the identification of the need for specific forms of remedial training, may be a stated purpose for measurement. Diagnosis may be defined as the classification of performance so that the resulting classes map 1:1 with potential remedial action. The classification may depend on any of the forms of performance measurement discussed above, and identification of the specific measures to be selected require that the measurement designer have some understanding of the underlying behavior (possibly determined through structured empirical tests).

There are two approaches to diagnosis which will be treated in a later section. One is a stochastic approach while the other uses expert system techniques derived from the field of artificial intelligence. The first has the advantage that the statistical methods help to pick a useable measure set from a larger set of candidate measures, while the second approach leads to specification of the measures through analysis of the knowledge of expert diagnosticians.

SECTION 7.

COMMENTS ON PERFORMANCE DIAGNOSIS FOR RESEARCH AND TRAINING

The preceding sections have discussed the structure of measurement, sampling considerations, segmentation logic and possible transforms. When selecting and developing measurement for research or training, there are performance diagnosis techniques which can be used to refine measures, develop global performance functions, and examine complex performance.

Two approaches to automated performance diagnosis will be discussed here. One is a statistical technique which uses performance data from known groups to derive algorithms for predicting group membership from test samples of performance data. The other is a technique adapted from artificial intelligence research which determines if test samples of performance match anomalous characteristics specified by subject matter experts.

MULTIVARIATE DISCRIMINANT ANALYSIS

The multivariate discriminant analysis produces a linear combination of available data as a single computed variable which best discriminates between the performance of groups with known characteristics. For example, these groups may consist of expert and novice performers, or groups of desired performance and specific performance inadequacies. The resulting discriminant functions can be used on data acquired during later training or research to obtain a classification to characterize the performance which was exhibited.

A geometric interpretation of discriminant analysis is shown in Figure 7 (from Cooley and Lohnes, 1971). The figure shows the case of two groups and two performance measures, but keep in mind that a much larger number of groups and measures is possible. Each circle is the locus of points of equal frequency for a group. The pairs of points for the intersection of corresponding contours in each group define a straight line (II), and we may also construct another line (I) perpendicular to line II. If the points in the two-dimensional space are projected onto line I, the overlap between the two groups will be smaller than for any other possible line. The discriminant function is a transformation which combines the individual performance measures into a single score, and that score is a location along line I. Of course, real human performance data are seldom as clear as this illustration.

The point b divides the discriminant line into two regions, one indicating probable membership in Group A and the other indicating probable membership in Group B. Suppose that Group A consists of data collected from subjects exhibiting desirable performance and that Group B is from subjects exhibiting a specific performance anomaly. The discriminant function can be used to classify the performance of a test subject to one of the Groups; that is, the discriminant performance can be used to diagnose the subject's performance. If the procedure was extended to

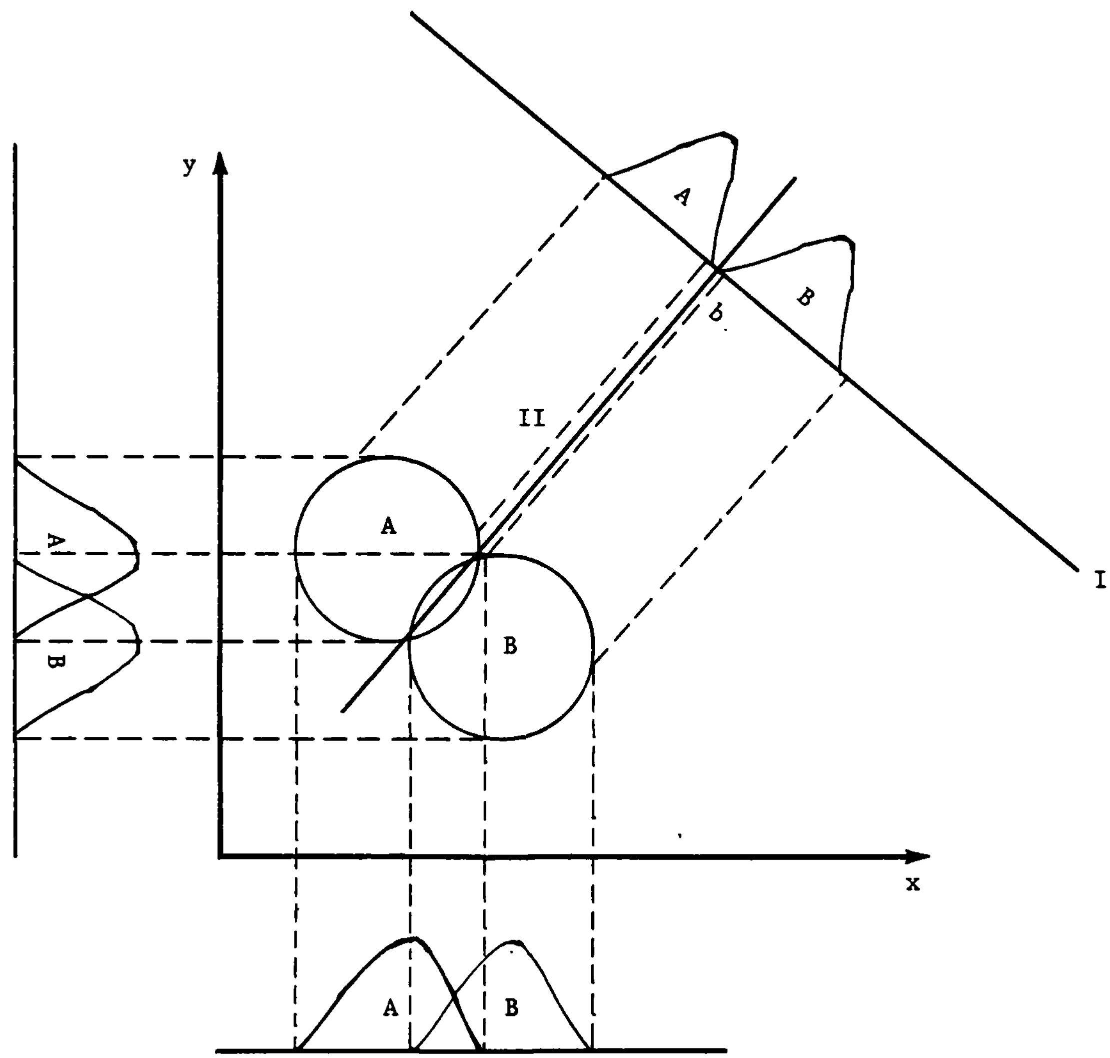


Figure 7. Geometric Interpretation of Multiple Discriminant Analyses.

include Groups C, D, ..., each with a specific performance anomaly, the diagnosis can indicate the identity of the performance deficiency and corrective action can be taken (e.g., remedial training).

Suppose that Group A consists of data from expert performers and that Group B consists of data from novices. The discriminant analysis can be conducted with a relatively large set of candidate measures, and it can be noted which of these contribute heavily to the discriminating power. Those which do not contribute to the discrimination can be eliminated as these measures can be viewed as being insensitive. Thus, the discriminant analysis can be used as a measure selection tool, even if the discriminant function itself is not directly useful. Common univariate analyses can be used for the same purpose, but correlations between measures might not be sensed.

Quantities of data might be required for fruitful application of this statistical technique. Data ordinarily are collected for any measures which are suspected of contributing to an ability to discriminate between the required classes of performance, and then the most useful set of measures are weeded out. Data must be collected for a representative set of subjects and for a number of repetitions large in comparison to the number of different types of performance measures. The end result is that extensive data collection might be required.

Over the years, we have developed a set of multivariate measurement analysis routines, some of which can reduce the amount of data which is required by standard multivariate analysis methods. These routines can be used to (a) remove highly correlated measures, (b) remove data "outliers" which tend to distort correlation based on least squares regression criteria, (c) adjust the resulting model for possible "overfit" and "shrinkage" of weighting coefficients by applying "ridge-regression" techniques to the multiple discriminant analysis, and (d) provide a non-parametric discriminant analysis based on Tukey's "Quick Test of Location" for those cases where there are insufficient observations or degrees of freedom to use a parametric model.

Fragments of these techniques have appeared in our past reports, but there is no source which has put together these methods and our adaptations of them in one document. We thought it useful, therefore, to include a discussion of these techniques as Appendix C. Also included are FORTRAN program listings of the analyses. The reader is referred to Appendix C for greater depth, and to assess the tradeoff between potential benefits and associated costs of data collection and analysis associated with these techniques.

KNOWLEDGE-BASED SYSTEMS APPROACH

Computers currently excel in high-speed calculating and in storing and retrieving large amounts of data; however, such computations can solve only a fraction of human problems. The challenge is to have machines that are capable of "thinking," or at least emulating the ways humans draw upon past experience to solve new problems. Work on computerized expert systems grew

out of earlier work in artificial intelligence to make computers perceive, reason and understand.

This work has now progressed out of the laboratory and there are over two dozen practical systems in operation. Knowledge-based systems perform some of the most difficult decision-making jobs that include use of judgment, rules of thumb, and experience. This approach is of particular interest here because there has been success applying these techniques to automating human decision ability where the "spot-a-pattern, draw-a-conclusion" style of reasoning is used by experts.

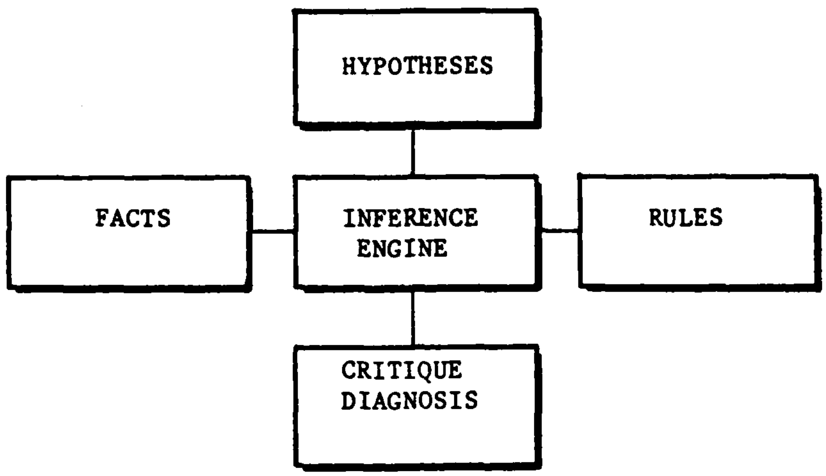


Figure 8. Block diagram of a knowledge-based system.

A specific form of knowledge-based system which can be applied to the task of human performance diagnosis is shown in Figure 8. It is often called a production rule system, rule-based system or if-then system. For the current application, the hypotheses will be specific diagnoses, the facts will be selected performance measures, and the rules will be if-then statements which collectively permit the knowledge-based system to conclude whether any of the hypotheses can be substantiated.

The Inference Engine is the control part which attempts to conclude that one of the hypotheses is true. A subtle key to the effectiveness of this approach lies in the separation of these parts of the knowledge-based system. Furthermore, the importance of these parts is not equal, however, as there is some consensus in the Artificial Intelligence community that the power mainly lies in the knowledge base (the rules), and not in the inference procedure.

Requirements for knowledge acquisition. An extensive, labor-intensive effort is required to extract subject matter expert knowledge required to define the knowledge base of rules. Knowledge is currently acquired in a very painstaking way in which individual specialists (Knowledge Engineers (KEs)) work with Subject Matter Experts (SMEs) to extract a complete and

consistent knowledge base. The information to be extracted from the SME includes not only the factual knowledge which may be found in textbooks, but also the rule-of-thumb heuristics which are not published and which the SME may not be able to articulate readily. While an abundance of data is not available, it appears that the knowledge acquisition process may have required 2-3 years for existing large-scale systems (e.g. medical diagnosis).

Inferencing methods. There are two basic methods for automated inferencing: the forward-chaining method and the backward-chaining method. The forward-chaining method simply tries all rules; if any new conclusions are derived, they are added to the stored facts, and all rules are tried again. This process stops only when a complete pass through the knowledge base yields no new facts.

The backward-chaining method starts with specific hypotheses and searches for the rules which collectively can conclude the specific hypothesis. A hypothesis is selected (in order on a list) and a rule is found which can conclude that the hypothesis is true. The if-part of this rule is examined to determine whether an if-statement is part of the Facts list, or whether there is another rule which can conclude that the if-statement is true. The latter involves scanning the rules for a rule with the appropriate then-part, and then attempting to determine whether the if-statements can be verified. This verification process is recursive, with the verification function using itself repeatedly.

While the two inferencing procedures are similar, there are some substantial differences. The forward-chaining method requires no statement of hypotheses, the procedure simply derives any conclusions which are possible, whether they are hypotheses or just intermediate facts. All rules are exercised repeatedly; therefore, the backward-chaining procedure may be quicker. The backward chaining procedure, however, requires large amounts of computer memory for processing, while the forward-chaining procedure requires little more than the static storage capability.

An example. In air combat maneuvering, the High Yo-Yo is a maneuver designed to be used at low angle-off (the aircraft axes are nearly in alignment), low aspect angles (the angle to the attacker off the defenders tail is small), and high overtaking speed. When attackers realize their present overtake will cause a flight path overshoot (which may make them a defender), they may use the High Yo-Yo, which is a maneuver out of the present plane of turning. This rule may be expressed as the following LISP statements (other AI languages could be used as well):

```
(IF (OFFENSIVE)(HIGH OVERTAKE)(ANGLEOFF < 40)) (THEN (USE HIGH YO-YO))
(IF (CLOSING RATE < -1500)) (THEN (HIGH OVERTAKE))
```

Furthermore, if the pilot executes a High Yo-Yo, there are a number of errors which occur frequently (Common Student Errors). These may be expressed as rules such as the following (TBD = to be determined):

```
(IF (HIGH YO-YO)(ASPECT > TBD)(ANGLEOFF > TBD))(THEN (COMMON STUDENT
ERROR 1))
```

(IF (HIGH YO-YO)(BEGIN HIGH YO-YO BEFORE TBD)(PITCH ANGLE > TBD))
(THEN (COMMON STUDENT ERROR 2)(EXPECT DEFENDER TO UNLOAD AND SEPARATE))
(IF (HIGH YO-YO)(BEGIN HIGH YO-YO AFTER TBD)(PITCH ANGLE < TBD))
(THEN (COMMON STUDENT ERROR 3)(EXPECT TO OVERSHOOT))
(IF (HIGH YO-YO)(DURING ROLLOUT)(G < TBD))(THEN (COMMON STUDENT ERROR 4)(EXPECT TO OVERSHOOT))

Note that this set of rules is only suggestive and not complete. However, it may be seen that the inference engine can conclude that the conditions are right for the pilot to execute a High Yo-Yo, and measurement can proceed to determine whether such a maneuver was executed.

This provides some capability for diagnosing the pilot's decision making. Further, given that High Yo-Yo is added to the Facts (being executed), the inference engine can proceed to further diagnose whether expected errors occur. During this period the High Yo-Yo Common Student Errors would be added to the Hypotheses, and backward chaining would seek to verify each of these common errors.

If the set of rules was no more extensive than defined above, there would be little need for artificial intelligence techniques. For a problem as complex as air combat maneuvering, however, one can establish hundreds of rules, requiring a more-sophisticated implementation. Nevertheless, once such a system has been developed and tested, it may be possible to reprogram in more conventional languages (e.g. FORTRAN or PASCAL, rather than LISP) and possibly with more conventional techniques.

COMPARISON OF APPROACHES

Diagnostic technique. While the same goal is shared by the multivariate and knowledge-based approaches, the basic source of diagnostic material is quite different. The multivariate statistical approach derives diagnostic algorithms from samples of pre-classified performance data; the knowledge-based systems approach derives inferences from rules extracted from the knowledge of subject matter experts. The multivariate statistical approach may have an advantage where the basis for diagnosis is obscure, unknown, or beyond the ability of subject matter experts to articulate. The knowledge-based systems approach may have an advantage where the rules of diagnosis may be easily extracted from existing knowledge.

Processing requirements. The multivariate statistical approach requires a computer with fast number-crunching and large-data base capability to perform the processing required to identify minimal sets of measurements and algorithms for diagnosis. Once this process has been performed, however, diagnosis can be performed with little additional programming and ordinary computer capabilities. On the other hand, the database of rules can be quite large and processing slow for implementations of the knowledge-based systems approach.

When backward-chaining recursive procedures are used, very large memory requirements can result. For example, given a microcomputer with 256KB of memory, it is possible to do backward chained inferencing with about 300-

400 rules, and to do forward chained inferencing with about 2000-2500 rules. Of course, with a full scale major implementation, artificial intelligence researchers have sought out the largest mainframe computers available.

As a further complication, programming is eased for this type of processing when LISP or PROLOG languages, or their derivatives, are used. Programs written in these languages are not mixed easily with programs written in common data processing languages such as FORTRAN or PASCAL. Diagnosis, however, does not have to take place in real time; this permits a two-pass approach to be taken, and standalone microcomputers can be interfaced to accomplish limited or focussed diagnosis.

Data collection. As pointed out above, the multivariate statistical approach requires extensive collection and analysis of performance data, while the knowledge-based systems approach requires extensive extraction of subject-matter-expert knowledge. Either task can be formidable; however, for major levels of diagnosis, it is currently believed that the knowledge-based systems approach can lead to a working product more quickly.

Allocation of intelligent functions. Diagnosis does not have to be automated completely to satisfy all requirements. In fact, in the short term, it is unlikely that computer diagnosis will completely replace human diagnostic functions. Certainly it is desirable in all cases to have the human diagnostician participate for a long period of development to refine the system to achieve human acceptance.

For some applications it is necessary only to augment human diagnosis--to ensure that performance does not go unobserved, and suggest a diagnosis. In some cases one may wish only to provide the human diagnostician with sensitive discriminating information. There is, then, a potential allocation of diagnostic functions between human and computer, and, probably not all applications will require fully automated diagnosis.

Comprehensive identification of errors. Both of the automated approaches which have been discussed require the identification of all performance anomalies of interest as a starting point. It is relatively easy to identify common student errors, because these are listed easily by experienced instructors. It may be difficult, however, to make a comprehensive list of all performance anomalies which could be encountered by a performance diagnosis system.

There are no error taxonomies readily available. Fault analysis for automatic systems often involves definition of a fault tree; for example, each significant failure is identified, then the events leading to these failures are identified, and then the events which could lead to these events are identified, and so forth. Clearly, this is a major undertaking for any but the simplest systems. Consequently, a fundamental and unresolved problem is the identification of a comprehensive taxonomy of performance anomalies; given this, there are techniques which offer some promise for automated or semi-automated performance diagnosis.

SUMMARY

Two methods have been described; each has potential for automated performance diagnosis. Multivariate discriminant analyses may be the method of choice where differences are obscure and experts are not able to articulate a basis for diagnosis. The knowledge-based systems approach may be the method to use where an emulation of expert inferences is desired. The two approaches are not mutually exclusive; they could be used in a complementary way, and conceivably a diagnostic system could be designed using both. In any case, there is emerging technology upon which a system for automated diagnoses of human performance could be implemented. As a consequence, future performance measurement systems should move to a new and higher level of meaning.

SECTION 8.

DISCUSSION

This section will present some of the complexities of real-world measurement to provide information which should be considered when applying the measurement methods covered in the foregoing sections. Measurement in the context of real flight missions must consider specific tasks, performance objectives, tradeoffs, and purposes for measurement; these are discussed in the initial paragraphs of this section. Then, measurement considerations imposed by different end purposes (system design, personnel selection, individual training, operational training, and behavioral research) are amplified. Real-time performance measurement for training systems is being automated; a discussion also is provided to reveal some of the issues associated with this trend.

FLIGHT PERFORMANCE MEASUREMENT DOMAIN

Consider Figure 9; it lists flight phases (global tasks), performance objectives and purposes of measurement. Within each phase there are tasks such as turn, climb, descend, accelerate, and decelerate relative to the airmass, another aircraft, the terrain or objects on the ground or water. Aircrew knowledge, cognitive, perceptual and motor abilities are embedded in the performance of these tasks, are inferred by task performance measures, and cannot be measured directly.

When performing flight tasks, pilots have performance objectives, which are classified as survival, effectiveness, efficiency and regulatory in Figure 9, but there are overlaps within these categories. Survival is paramount; without it there is no effectiveness, unless much more than the survival of the individual aircraft is at stake. Without effectiveness, efficiency is meaningless. Regulations take precedence over effectiveness and efficiency in peacetime, but in wartime the relative importance of many objectives change. It might be important to learn how these performance objectives might change in war.

The flight situation at every moment will affect the way experienced pilots trade-off which objectives are optimized, and how much error is acceptable, given the existing workload, current and projected states of the aircraft. For example, during take-offs, climbs, descents, approaches, landings, instruments and emergency procedures, one can measure error from the required profile in each degree-of-freedom (such as pitch, roll, heading, altitude and so forth) and profile dimension, as illustrated earlier in this report. But pilots control these dimensions simultaneously and make trade-offs; also, error from any dimension, or accuracy, is only one of several potential measures of effectiveness and efficiency.

Moreover, as one addresses flight tasks and missions of operational complexity, the performance objectives of the pilot can change. During UPT, pilots are learning the basics, and are expected to fly each maneuver as precisely as possible at all times. After training, pilots are expected

to apply judgment, and overattention to accuracy of flight at the wrong times might compromise safety and mission performance.

Pilots learn when to attend to each task, and the degree of precision needed for each flight situation. For example, in heavy turbulence, speed and altitude may be permitted to vary more than usual to prevent airframe damage. When navigating, a pilot who is slightly off course may not return to the course centerline, but may make a slight heading change to cross near the next checkpoint in order to optimize fuel consumption; a measure of error from the course centerline would not provide adequate performance information. Perfect profiles might mean that pilots are spending too much time looking in the cockpit, and not enough time outside.

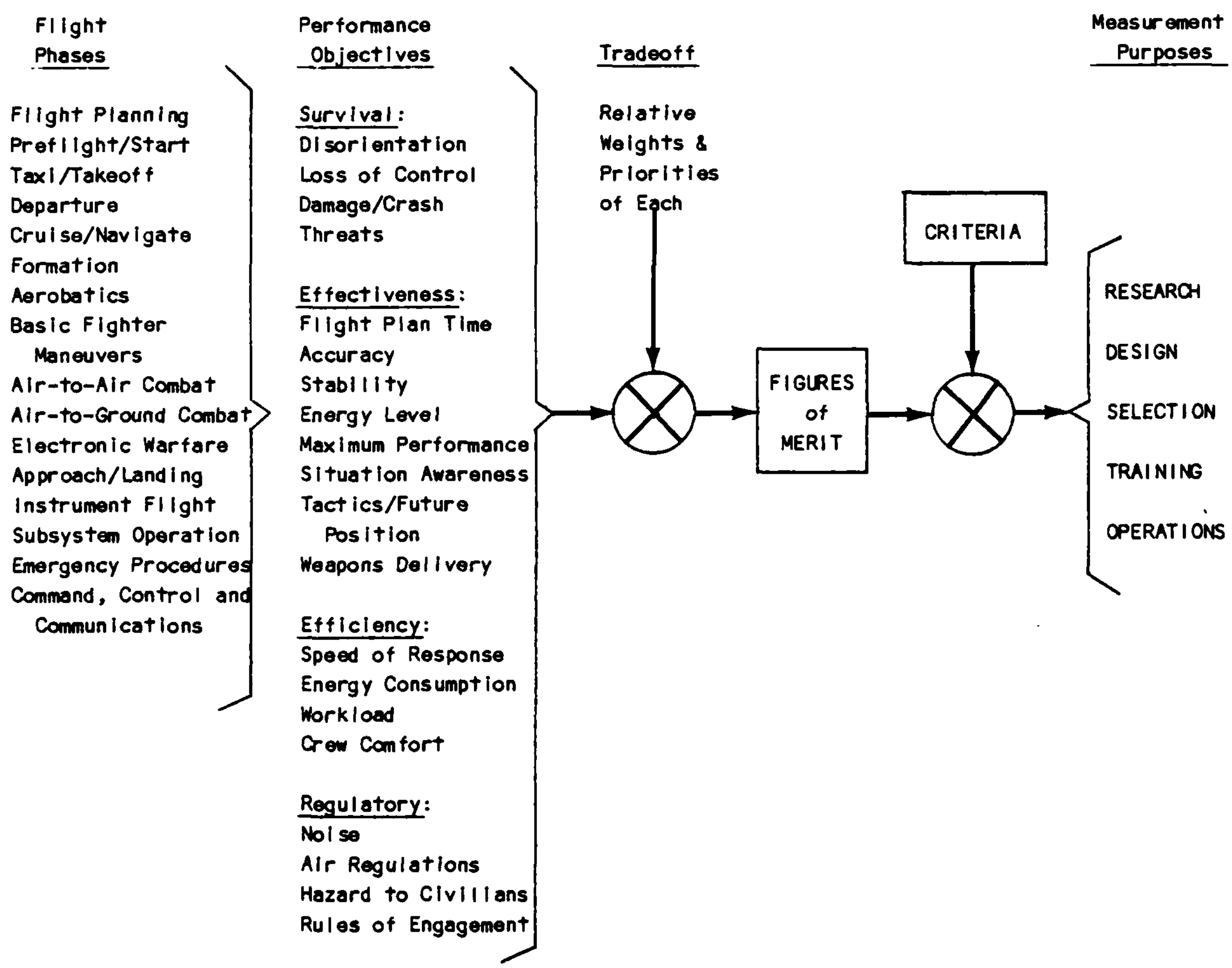


Figure 9. Flight Performance Measurement Domain.

As training progresses from UPT to operational readiness, pilots become more skilled, and onboard systems increase in sophistication, changing the nature of the job. Performance objectives change and measures should reflect these changes; it may no longer be appropriate to measure heading, height and airspeed the same way. Tactics, energy management, coordination with other aircraft and ground forces, and decision making become important tasks.

There is a mapping of measures from flight tasks through performance objectives. If it were constructed, this map would show the relative priorities and weights of each measure on a combined figure of merit for each task. One would expect these priorities and weights to change with pilot experience, the nature of the flight environment, and the mission requirements.

For each purpose of measurement in Figure 9, criteria for acceptable and unacceptable performance are needed to evaluate performance for each task. Information for research, system design, personnel selection, and training may require slightly different measures and criteria than for operations, but criteria should be derived from operations because the ultimate goal of the aircrew is to perform its operational purpose in both peace and war.

For some purposes, such as research on basic human perceptual, motor and cognitive abilities, or training diagnosis, a combined figure of merit may be too insensitive to provide needed information; exclusive use of a single figure of merit is not recommended, but the measurement user should know the relationship between individual measures, overall figures of merit and operational criteria to assess the importance of performance changes to real world operations.

Commonly stated criteria (such as hold heading, altitude and airspeed within certain limits) for aircrew task performance are general performance boundaries. They have emerged to guide instructors and flight inspectors who observe much more of a pilot's behavior than these criteria suggest; taken alone, such criteria may not provide sufficient information. They tend to obscure the relationships which are illustrated in Figure 9, and ignore interactions between dimensions of performance (and of measurement), all of which are assessed by the human instructor, whose judgments are based on knowledge and years of experience which is difficult to extract and quantify.

Where there are no fixed profiles, such as during air-to-air combat, only terminal parts of the task, such as gun tracking or outcome, usually are scored. During introductory training, the Fighter Weapons School at Nellis AFB creates set-ups in such a way that a tracking solution can be achieved ONLY if the pilot performs properly. The ACM ranges measure everything that happens, and makes these data available for replay, but only kills are scored.

During ground attack weapons delivery training, it is common to score only the weapons impact. Terminal measures do not provide information on

how the pilot maneuvered to achieve the kill or drop the bomb. Diagnostic information is missing, although an error analysis of the miss distance and direction can provide some guidance for the pilot who remembers the exact sight picture and aircraft states at the time of release. The accuracy of perception and memory for precise details during such extreme workload conditions as the final seconds of air-to-ground weapon delivery can be questioned.

Again, the judgments of expert observers are used to provide necessary guidance for pilots, and assess performance in current practice; the basis for those judgments has to be captured to develop meaningful measures and criteria.

What, then, are our measures telling us about the knowledge, cognitive, perceptual and motor abilities which cannot be directly measured? What are measures derived from current practice and existing criteria telling us about behavior, if they may not reflect the real behavior of experts? Is zero error always the best? Most measurement in research and training seems to make this assumption. Are we measuring properly for system design, personnel selection and training, and operational readiness assessment? The current state-of-measurement for each of these purposes is discussed next.

SOME CURRENT MEASUREMENT ISSUES

A overview of performance measurement issues in (a) system design studies, (b) personnel selection, (c) individual training, (d) operational training, and (e) behavioral research follows:

System Design Studies. The purpose of measurement during system design is to provide information which will predict the capabilities of aircrew members in aircraft systems, and to provide this information in a way that relates to system design questions early enough in the design cycle to have an influence on it. Devices used to study system design issues include static mockups, full scale mission simulators, and sometimes aircraft. Because system design alternatives may cover a range of configurations and capabilities, the most informative data would be validated performance models of the human operator which could be used with aircraft and weapon system models to predict performance in computer simulations.

The human aspects of aircraft and weapon system designs usually are concerned with (a) operator station design and use throughout the full range of the environment in which the system will have to perform, (b) the precision of control and stability, (c) the level of automation that will be required to assist human control for precision and safety, (d) operator workload, cognitive, perceptual and psychomotor functions, (e) reductions in operator capacity due to physical, environmental, radiation or workload stress, (f) operator potential error rates, and (g) personnel and training requirements. Each of these purposes requires slightly different measures.

Measurement may be derived from manual observational data, automated instrumentation and measurement systems, and human operator and system

models. We can measure response time, switch positioning accuracy, and ability to reach controls and read displays. We know how to model and measure operator dynamic response characteristics (and operator-system precision and stability) in some situations (cf. McRuer and Krendal, 1957; Muckler and Obermayer, 1964; Baron, 1981; McDonnell Douglas, 1981), but not all of them. There are workload assessment measures and models, but we have not come to any agreement on valid measures of mental workload or perceptual capacity, let alone establishing optimum levels for design use.

We know how to measure error from predetermined profiles, but we don't necessarily know how to scale performance quality where there may be many solutions to the problem in changing or emergent environments. Even though we know how to measure many tasks individually, we do not necessarily know how to evaluate the importance of each measure to overall performance, except through the use of empirical data analysis and modeling techiques.

Often, system performance measures have to be used to infer correct interpretation of all displays and communications, and these measures will be affected by the way in which the test scenario is constructed. Since performance will be a function of what the aircrew is being asked to do, measures will be related directly to the parameters of the test. Neither the measures that are taken nor the test conditions for measurement are standardized.

Where safety (both of the system and the environment or population it might affect) is important, it is of special concern to acquire valid human error data. In spite of renewed interest in this area by the Nuclear Regulatory Agency, there is a lack of valid and generalizeable human error data for system design purposes. Human error rates can be sampled in a simulator for design purposes, but without knowledge of the error rates in the real world, the representativeness of the sample can be questioned.

Furthermore, it is unlikely in any system design program that enough time or resources can be allocated to collect the amount of data needed to predict operator error rates with confidence, especially for the low probability but high cost events. And, there are serious issues concerning "what constitutes a test" for such events. What can replace the hours and perhaps years of training and everyday operations that have preceded major aircraft accidents?

A related issue is: what constitutes a test for aircrew members who are exposed to extreme heat, g-loading or toxic environments, and how are performance measures to be related to mission effectiveness criteria? How much performance degradation is acceptable under what circumstances?

It is clear that aircraft and weapon systems are designed and delivered, and they work, although many systems require operational test, evaluation and field changes. They work because of the experience of designers, analysts, and subject matter experts, who often do not have adequate performance data, models and criteria. In the absence of data, design decisions can be wrong, and systems can emerge with less than optimum demands on the human operator, which can have an adverse affect on

the training requirements and mission effectiveness. Valid performance data, models and criteria can reduce these risks.

Personnel Selection. Personnel selection criteria usually are derived from paper-and-pencil tests, review of academic credentials, interviews and medical examinations. Paper and pencil tests predict academic performance, but do not predict job performance very well. There are questions about basic knowledge, abilities and human performance capacities (as well as social and attitudinal factors) that must be part of selection tests. All three military services have extensive programs to improve the ability of selection tests to predict job performance.

The use of simulators and part-task devices to aid selection decisions by measuring performance on job-sample tests has been examined (Long and Varney, 1975; Shipley, 1983; Kozinsky and Pack, 1982). Since job sample tests have to be designed so that a minimum amount of training is needed, measures similar to those used early in training should be useful. It is interesting to note that Shipley (1983) found noticeable differences in the ability of selectees to handle transitions from one steady state to the next.

The measures for this purpose require longitudinal validation, and without operational figures of merit and criteria, the best that can be done is to validate selection measures against success in training or job performance batteries. More work in this area is needed.

Individual Training. There has been great improvement in training performance measurement brought about by Instructional System Development (ISD) efforts throughout the military services. The ISD model requires the development of specific behavioral objectives for each task to be trained, and specification of performance standards. For those tasks in which specific procedures are to be followed, and for which performance standards can be specified, measures based on "school solutions" are within the state-of-the-art.

ISD methods, however, may be insufficient for specifying measures for tasks which involve extensive maneuvering or are reactive. In current practice, maneuvering tasks often are judged best by expert observers, because the decision rules are complex. When attempting to reduce these rules to printed procedures, "the book" can be wrong, as was found by Knoop and Welde (1973). An examination of a small sample of ACM engagements in the Simulator for Air-To-Air Combat by Wooldridge and Obermayer (1982) found that published rules of thumb for ACM were not followed by pilots who were judged to be experts. Measurement which is based on rules which are published as initial training material may not be sufficient.

Operational Training. Aircrew members receive continuation training and operational "readiness" evaluation. Numeric performance criteria are applied to a few tasks, such as the annual instrument flight check and the number of qualifying bombs for ground attack. If one is lucky enough to get time on an instrumented range, ACM performance is recorded for playback and kills are scored. But, performance is not measured for many tasks.

Proficiency is presumed by "square filling" exercises, such as the number of instrument approaches, number of low level flight mission, number of night landings, number of times one entered the ACM practice areas, and so forth. These are measures of experience, but they are not performance measures or criteria. Consequently, there are no current data upon which to build figures of merit and performance criteria for presently unmeasured or assessed tasks. Without such criteria, how to we train individuals to be experts?

In most operational aircrew tasks, individual crew members work as a team, which reacts to the environment which may be changing and interactive with them. If the environment changes, the situation may be "emergent," and have no fixed solution which is best all of the time (as opposed to the school solution).

Airspace control is an example of a cooperative emergent environment, where the controlled aircraft are working with the system for the safety of everyone. Military combat, and battle command and control exercises are examples of uncooperative emergent environments, where flight crews must react to adversaries as well as coordinate with their own forces.

In military exercises, unit performance is measured usually at the outcome level, such as the resources used, ordnance expended, number of casualties inflicted or targets destroyed, number of casualties sustained, and the time to complete the drill. While these data are useful, seldom does one find measures which can diagnose the cause of a particular outcome in a way that can prescribe directly what training is needed by individuals to improve crew performance.

Individual crew members may "drop the ball," but their errors might be mitigated by another crew member, another flight element, a good forward air controller, or an error by the adversary. Conversely, an individual crew member may not perform well because someone else did not do their job properly. Measures of total crew performance do not necessarily provide the performance data and feedback for individual crew members.

Measures of individual and crew performance in emergent situations have not been well developed, and criteria are lacking. The usual solution, in systems that are so equipped, is to setup a scenario, record performance and communications in time history form, then replay the exercise to debrief the players. Some systems provide real-time scoring of casualties and feedback for individual players (e.g. air combat maneuvering ranges), but if the scenario is changed, the outcomes are likely to change. The assessment of the quality of performance is scenario specific.

Individual and system performance models, which account for all the dynamic elements in the scenario that influence individual decisions and system performance, appear to offer the only solution to performance measurement and assessment in emergent situations. Without such models, the effects of individuals on team performance, and extrapolation of team performance from one situation (scenario) to another would be difficult to quantify. But, there are many unresolved issues related to the structure,

content, and fidelity of such models; the design and development of models and measures for individual and team performance assessment will require research and empirical testing to determine their validity.

Thus, the measures which are taken in operational situations today contain very little information on which to assess the maintenance of skill, insure individual proficiency or the proficiency of units, or expose the real state of mission readiness. In the few cases where there are performance criteria, current measurement practice is too imprecise to provide the kind of data and criteria which are required for improving system design, and measurement system design for personnel selection and training beyond current practice.

Behavioral Research. All of the above measurement issues apply to research and development of human-machine systems, for design, personnel selection and training--and moreso. More measurement precision is required for research than for "everyday" training or operations, to search for previously unknown knowledge, to reduce experimental error, and to predict performance for a variety of purposes.

The more measurement captures relevant human knowledge, abilities, learning, performance, individual differences, and the varying nature of dynamic tasks, the less unaccounted for variance will appear in the data. As the performance data space precision in the experiment is improved, so will be improved our ability to predict performance in the real world.

As stated before, there are insufficient operational criteria and figures of merit to adequately support research. Many measures typically are taken in research studies, but often the relationship between the research result and real world mission performance is not known. There have been attempts to relate research performance data to operationally relevant criteria (Westra, Simon, Collyer and Chambers, 1982; Brictson, Burger and Wulfeck, 1973; Knoop and Welde, 1973); but, in general, it is difficult to map performance data from research studies into overall system or mission effectiveness, because these criteria are unknown.

Even when transfer-of-training studies have been permitted, the data usually show only the "replacement value" (in terms of transfer effects, training effectiveness or cost effectiveness) of the whole training device within the existing training curriculum. If the device contains features that train additional skills, and the curriculum is not adjusted to the change in training content which is brought about by the new device, a transfer-of-training study will not show the true value of the device.

Also, current transfer-of-training measures cannot provide information on the elements or features of instruction that are transferring positively and negatively. Consequently, use of the transfer effectiveness ratio as the only metric of training value has been questioned (Rolfe and Caro, 1982). New test and measurement methods are needed; we may have to examine transfer on a task by task performance measurement basis.

The analysis of behavior for each particular research purpose requires the development of (a) measures which are sensitive to the performance

changes that occur, (b) criteria for selecting and combining measures to develop overall figures of merit, and (c) criteria to relate those performance changes to operationally relevant factors. Since we lack detailed measures of "on the job performance" for many tasks, behavioral research suffers the criterion and measurement validity problems.

DESIGN ISSUES IN AUTOMATED MEASUREMENT SYSTEMS

Automated measurement systems are being specified as an integral part of the requirements for many new training systems. When designing a real-time automated measurement system, several issues emerge. Automated measurement systems remove some of the judgment and insights of human instructors from the measurement role. As discussed earlier, some of the intelligence of human instructors and examiners has to be defined well enough to make the systems work. There are six areas of challenge:

Knowledge of Tasks. The measurement system must know what tasks are to be performed; this is easy for fixed profiles or procedures, but it can require pattern recognition and probability assessment if the task changes as a function of the simulated environment, or there is no exact profile or procedure.

Segmentation. The system must recognize the start of a task and the end of it; the measure segmentation rules discussed in Section 5 represent initial solutions only; all aircrew tasks do not start or end with events which are recognized easily. Some tasks are performed in parallel. Often, there is a flow from one task to the next, where the change from any task to the next may be subject to the way the aircrew chooses to do the job, or may be subject to a variety of environmental events or factors which are external to the specific task at the moment. Expert human observers usually take all of these factors into account; automated systems will do so only if they are programmed to do so, and some tasks, such as monitoring and communications, may not be measured directly.

Performance Diagnosis. Performance diagnosis requires identification of all possible errors and construction of measures for many of them; this would involve much more analysis than is common in most ISD efforts, and the diagnosis of prime causes is non-trivial. The diagnosis difficulty lies at a level below the obvious blunder, where minor deviations from expected performance compound into an error. Here, measurement must be able to spot patterns over time, assess probabilities and determine probable causes. Examination of time histories can lead to after the fact diagnosis, but real or near-real time diagnosis could require significant processing resources, and extreme care must be taken to capture only errors and not issue false alarms.

If the purpose of diagnosis is to guide training, decision algorithms may require examination of more than the current exercise. Information on past student performance, the rate of learning and future exposure to the same training event might be needed to determine the best course of action. If the student knew what to do, perhaps he or she just needed practice; if the opportunity for more practice will be presented in the future, no

action may be necessary. If no further exposure to the task is provided, additional practice may be required. If the student did not know what to do, either academic or simulator remediation may be required. These are training strategy issues, but they impact on performance measurement system design for real or near-real time application.

Summarized performance metrics are not likely to provide diagnostic information. To develop the required information, all factors which affect task performance will have to be known. Computation, memory and storage resources will have to be dedicated for this purpose. The architecture for storage, retrieval and processing of diagnostic data may be different from that of the basic simulation; LISP, PROLOG or relational database methods may be the best way to organize and access these data.

Often, there are limited resources that can be provided for measurement purposes, so the value of the information has to be traded-off against the cost to provide it. There are no general methods to determine the training value of various levels of diagnostic information. Guidelines for the presentation of required information for instructional use and diagnosis for real or near-real time measurement systems in simulators have never been studied or formally developed.

Additional Measures. System performance measures may not be sufficient for diagnosis or assessment purposes. Where learning has been measured for basic psychomotor and cognitive tasks (such as video games), there have been enormous individual differences in both performance level and learning rate at every stage. Contrary to popular belief, performance on relatively simple psychomotor tasks does not asymptote as early as many researchers and training specialists presume (Kennedy and Bittner, 1977). People get better and better over time, and they do so at different skill acquisition rates; more than a few slow learners eventually perform tasks better than fast learners. If we measure performance and predict success before the rate of skill acquisition is stabilized, our predictions may be of questionable value; measures related to learning may be needed.

User Interface. Data formats have to be developed to communicate measurement information to each user in their language (an RMS error score may not be understood by the user). Even researchers can be overwhelmed by scores of numbers. Moreover, if simple performance measures are not sufficient for assessment or diagnosis, methods to provide users with the meaning of more complex data and information structures will be needed. Dedicated analysis efforts are needed to find the best way to communicate measurement information to the user.

Validity of Measures. As said earlier, system performance measures are used to infer knowledge and abilities. Performance measures taken early in training may or may not predict end of course performance, and they may not capture all of the behavior that is sensed by an expert observer. Measures of performance are indirect for many purposes; indirect measures should be validated.

There have been isolated studies of measurement validity for aircrew tasks (cf. Connelly, Schuler and Knoop, 1969; Vreuls and Obermayer, 1971;

Knoop and Welde, 1973; Brictson, Burger and Wulfeck, 1973; Waag, Eddowes, Fuller and Fuller, 1975; Vreuls, Wooldridge, Obermayer, Johnson, Norman and Goldstein, 1976; Wooldridge, Kelly, Obermayer, Vreuls, Nelson and Norman, 1982). These studies show it is possible to construct valid measurement. Generally, the measure sets are more extensive and precise than grading criteria as it is commonly expressed by instructor grading forms, but high correlations with instructor judgments have been reported.

Several of these studies have found that all measures of a task cannot be weighted equally. Some measures carry heavier weights than others; if a multivariate analysis cannot be used to establish weighting coefficients, the next best procedure is to use norms to scale all measures relative to their distributions, then add the z-scores to derive overall performance metrics. One study of ACM performance found a multiple discriminant model which accounted for variance and could predict group membership (and be related to outcome), but would be difficult for an instructor to interpret, except as an index of "how goes the engagement."

Another finding was that the measures and weights which best controlled automated training were not predicted by analytical methods; when they were derived by empirical data analysis, a 40% reduction in time to train to the same performance criteria was found on subsequent tests. The measure set which produced superior automated training contained control input and other measures which tended to reflect technique--behavior that an instructor would sense, but for which there are no numeric criteria.

In their excellent discussion, Waag and Knoop (1977) indicated several kinds of validity tests which are needed to develop measures for training. Content validity is established first, then there are three empirical tests of increasing stringency: One test examines the ability of measures to discriminate between opposite ends of the skill continuum. A second test examines the functional relationships with concurrent measures, such as instructor ratings. A third test examines the functional relationships between performance measures and variables such as time in training. A fourth test, which was not discussed, might test the relationships between performance, measures of learning, and performance on the job.

Waag and Knoop (1977) noted that the problem of validation, already complicated by the lack of a single, necessary and sufficient test, is made more difficult by the a lack of standardization of validation criteria for any one type of test. Suffice it to say that if automated measurement systems are to be used for training, the measures must provide information which is proper and shown to be valid for that purpose.

In summary, modern technology provides a basis for improved measurement information. There is little doubt that machines can measure performance more accurately than human observers in many situations, but real-time automated measurement systems may have to capture some elements of behavior that presently are being observed and assessed by expert observers without quantitative criteria. Improved front-end analysis techniques are needed. Also, the costs of analysis and automated measurement systems may or may not be justified. There are no known guidelines for determining when and where automated measurement might be of value, and where it might not.

SECTION 9.

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

Each purpose of measurement requires measures that provide information which often is specific to that use. There are some unresolved issues in the mechanics of measurement and degree of intelligence needed for real-time automated performance measurement systems. Methods and guidelines for cost-benefit analysis of automated measurement systems for for training are needed.

Criteria for selecting appropriate measures and combining them for overall assessment are lacking. Human-system modeling techniques and empirical measurement development methods to derive, develop and validate measures for many tasks are available, but required data collection efforts have not been funded to the extent necessary for development of criteria and validated measures for system design, personnel selection and training, and operational performance assessment. In general, operational figures of merit and criteria are undeveloped and unknown for many tasks.

In personnel selection, measures of job performance by qualified aircrews would help to quantify the predictive validity of selection tests and measures. In training, quantitative measures of how various flight skills are acquired would assist training system research and development, as well as provide information of potential benefit for the training system user. Criteria for the performance of experts is lacking; if developed, they would aid training of expert performance. Measures of individual contributions to crew performance are needed to provide proper feedback to individual team members. Valid measures of performance in emergent environments are needed.

In spite of some of these difficulties, progress has been made on real-time performance measurement systems. The measure segmentation methods which are suggested in this report have the ability to pin-point the start and stop of maneuvering tasks with here-to-fore unrealized accuracy. It is now possible to measure transition performance in ways that have not been possible in the past. Transforms which reflect the performance strategies and objectives used by experienced pilots have been suggested as potential alternatives to measuring only error from a known profile.

Certainly, the suggested methods can be used for flight profiles which are well defined, and where performance standards are explicit, as in much of undergraduate pilot and initial crew training in flight simulators. There are issues of combining measures into overall performance metrics which have to be resolved, but there are many acceptable methods for doing this, including the use of norms. There are issues of the presentation of measurement information to the user which have to be addressed as well, but for many standard measures, the user interface problem can be solved by attention to the issue in system design and development.

Issues such as the predictive validity of measures during training and capturing more of the knowledge and experience of instructors will require more research, and possible use of different data structures, such as those provided by the methods of artificial intelligence. Until we are sure that real-time measurement systems can capture more of the assessment capability of expert observers than now appears to be the case, automated measurement systems will have to be advisory in nature; instructors will have to review the re ults and have the capability to override any automated score.

For certain training and research purposes, the use of observational data collection should not be forgotten in the pursuit of automation. Expert and trained observers have the capability to sense many things that automated systems cannot, such as communications content, visual lookout doctrine, situation awareness, monitoring, and subtle cues that indicate a crew member is anticipating a future event. Observers have limitations, however, so research on the best use of observers and observational data should continue.

RECOMMENDATIONS

It is recommended that:

1. Measurement should be segmented using the methods and guidelines which are suggested. This will require development of a short-term memory window of relevant system states, appropriate regions of the phase plane, and flags which are set by logic to mark the time and occurrence of specific conditions.

2. Measures of transitions from one steady-state to the next should be used as well as steady-state measures which have dominated past performance measure sets. Transitions contain valuable performance information which is otherwise lost.

3. Transforms which measure pilot performance objectives and strategies should be used wherever possible. Some of expert pilot performance is not captured by measures of error from course centerlines, desired altitudes, airspeeds or headings.

4. Transforms which capture control technique in the frequency domain should be used. Such measures contribute to the training performance evaluation, and can provide data for performance models.

5. Human-system performance models should be improved for system design and development efforts. Models which reflect human error rates and performance degradation due to adverse environments would be especially useful. Such models also would be useful for the evaluation of alternative tactics and mission effectiveness studies.

6. Research should be conducted on methods to evaluate the performance of experts on complex, mission oriented tasks. Figures of merit which reflect operational performance and criteria are needed for system design and training. Methods to extract complex rules which are used by expert evaluators are needed to quantify the basis for their judgments, and develop knowledge representation techniques which can be used for real-time performance diagnosis.

7. Research should be undertaken to identify and develop methods for the diagnosis of performance by real-time measurement systems.

8. Research should be conducted on measures of learning and individual differences to be used in conjunction with performance measures for the prediction of end-of-course performance. Valid methods to identify potential failures early in training could save substantial training costs, and improved measures which are related to learning and training time might improve training efficiency.

9. The allocation of measurement function between automated systems and human observers should be studied. Cost-effectiveness and training benefit studies are need to learn cost-benefit functions for various levels of automated measurement systems versus manual observational measurement.

10. Future efforts to develop guidelines for performance measurement should analyze both user information requirements and tasks, since measurement is information for a specific purpose.

11. The instructor or expert evaluator should retain the prime performance assessment role until automated performance measurement systems have been demonstrated to be appropriate and valid for this purpose.

12. Research on observational data collection methods should be continued for those tasks and measurement environments in which automated measurement is not possible.

REFERENCES

Alluisi, E. A. Methodology in the Use of Synthetic Tasks to Assess Complex Performance. Human Factors, 1967, 9, 375-384.

Baron, S. An Optimal Control Model Analysis of Data from a Simulated Hover Task. NAVTRAEQUIPCEN 80-C-0055-1. Orlando, FL: Naval Training Equipment Center, 1981.

Brictson, C. A., Burger, W. J. and Wulfeck, J. Validation and Application of a Carrier Landing Performance Score: The LPS. La Jolla, CA: Dunlap and Associates, March, 1973.

Blackman, R.B. and Tukey, J.W. The Measurement of Power Spectra. New York: Dover Publications, Inc., 1958.

Connelly, E. M., Schuler, A. R. and Knoop, P. A. Study of Adaptive Mathematical Models for Deriving Automated Pilot Performance Measurement Techniques. AFHRL-TR-69-7, Vol I. Wright-Patterson AFB, OH: Air Force Human Resources Laboratory, October, 1969.

Connelly, E.M. and Zeskind, R.M. An approach to continuous evaluation of manual control system performance. Paper presented at the 19th Annual Meeting of the Human Factors Society, Dallas, Texas, October 1975.

Connelly, E.M. Manned System Performance as a Function of Display Characteristics. OTR-62-77-1. Vienna, VA.: Omnemii, Inc., June 1977.

Cooley, W. W. and Lohnes, P. R. Multivariate Data Analysis. New York: John Wiley, 1971.

Department of Defense. Review of Tactical Jet Operational Readiness Training. (U). Washington, DC: Office of the Secretary of Defense, November, 1968.

Hennessy, R. T., Hockenberger, R. L., Barneby, S. G. and Vreuls, D. Design Requirements for an Automated Performance Measurement and Grading System for the UH-1 Flight Simulator. Technical Report 1220. Alexandria, VA: U.S. ARMY Research Institute for the Behavioral and Social Sciences, June 1979.

Hennessy, R. T., Sullivan, D. J., and Cooles, H. D. Critical Research Issues and Visual System Requirements for a V/STOL Training Research Simulator. NAVTRAEQUIPCEN 78-C-0076-1. Orlando, FL: Naval Training Equipment Center, October 1980.

Kennedy, R. S. and Bittner, A. C. Development of a Navy Performance Evaluation Test for Environmental Research (PETER). In Pope, L. T. and Meister, D. (Ed) Symposium Proceedings: Productivity Enhancement: Personnel Performance Assessment in Navy Systems. San Diego, CA: Navy Personnel Research and Development Center, October, 1977.

Knoop, P. A. and Welde, W. E. Automated Pilot Performance Assessment in The T-37: A Feasibility Study. AFHRL TR-72-6. Wright-Patterson AFB, OH: Air Force Human Resources Laboratory, April 1973.

Kozinsky, E. J. and Pack, R. W. Performance Measurement for Training Simulators. Final Report NP-783. Palo Alto, CA: Electric Power Research Institute, 1982.

Long, G. E. and Varney, N. C. Automated Pilot Aptitude Measurement System. AFHRL-TR-75-78. Brooks AFB, TX: Air Force Human Resources Laboratory, 1975.

McDonnell Douglas Corporation. Human Performance Models. IR&D Report MDC-Q0860-2, Volume I. St. Louis, MO: McDonnell Douglas Astronautics Company, 1981.

McRuer, D.T. and Krendel, E.S. Dynamic Response of Human Operators. WADC TR 56-524. Dayton, Ohio: USAF, 1957.

Mixon, T. R. and Moroney, W. F. An Annotated Bibiography of Objective Pilot Performance Measures. NAVTRAEQUIPCEN IH-330. Orlando, FL: Naval Training Equipment Center, January, 1982.

Muckler, F. A. and Obermayer, R. W. The Use of Man in Booster Guidance and Control. Report No. NASA CR-81. National Aeronautics and Space Administration, July 1964.

Naval Air Systems Command. NATOPS Flight Manual, Navy Model AV-8A Aircraft. Washington, DC: Naval Air Systems Command, 1975.

Obermayer, R.W. and Muckler, F.A. On the Inverse Optimal Control Problem in Manual Control Systems. NASA CR-208. Washington, D.C., April 1965.

Pruitt, V. R. and Moroney, W. F. Energy Maneuverability Displays for Air Combat Training. Paper 0148-7191/80/1005-1185. Society of Automative Engineers, Inc., 1980.

Ringland, R. F., Craig, S. J., and Clement, W. F. Survey of Piloting Factors in Fixed-Wing, V/STOL Aircraft Design. NWC TP 5941. China Lake, CA: Naval Weapons Center, February 1977.

Rolfe, J. M. and Caro, P. W. Determining the training effectiveness of flight simulators: Some basic issues and practical developments. Applied Ergonomics, 13, 1982, 243-250.

Pope, L. T. and Meister, D. (Ed) Symposium Proceedings: Productivity Enhancement: Personnel Performance Assessment in Navy Systems. San Diego, CA: Navy Personnel Research and Development Center, October, 1977.

Shipley, B. D. Complexity of Flight Path Data as an Index of Skill in Piloting Performances from a Simulator Based Job-Sample Test. Fort Rucker, AL: U.S. Army Research Institute Field Unit, 1983.

Smode, A. F. and Meyer, D. E. Research Data and Information Relevant to Pilot Training. Volume I. General Features of Air Force Pilot Training and Some Research Issues. AMRL TR-66-99, Vol. I. Wright-Patterson AFB, OH: Air Force Aerospace Medical Research Labortories, July, 1966.

Vreuls, D. and Wooldridge, A. L. Aircrew Performance Measurement. In Pope, L. T. and Meister, D. (Ed) Symposium Proceedings: Productivity Enhancement: Personnel Performance Assessment in Navy Systems. San Diego, CA: Navy Personnel Research and Development Center, October, 1977.

Vreuls, D., Wooldridge, A. L., Obermayer, R. W., Johnson, R. M., Norman, D. A., Goldstein, I. Development and Evaluation of Trainee Performance Measurement in an Automated Instrument Flight Maneuvers Trainer. NAVTRAEQUIPCEN 74-C-0063-1. Orlando, Florida: U. S. Naval Training Equipment Center, 1976.

Waag, W. L., Eddowes, E. E., Fuller, J. H., Fuller, R. R. ASUPT Automated Objective Performance Measurement System. AFHRL TR-75-3. Williams AFB, AZ: Air Force Human Resources Laboratory, 1975.

Waag, W. L. and Knoop, P. A. Planning for Aircrew Performance Measurement R&D. In Pope, L. T. and Meister, D. (Ed) Symposium Proceedings: Productivity Enhancement: Personnel Performance Assessment in Navy Systems. San Diego, CA: Navy Personnel Research and Development Center, October, 1977.

Webster's New Collegiate Dictionary. Springfield, Mass.: G. & C. Merrim Co., Publishers, 1956.

Westra, D. P., Simon, C. W., Collyer, S. C., and Chambers, W. S. Simulator Design Features for Carrier Landing I. Performance Experiments. NAV-TRAEQUIPCEN 78-C-0060-7. Orlando, FL: Naval Training Equipment Center, September 1982.

Wooldridge, L., Kelly, M. J., Obermayer, R. W., Vreuls, D., Nelson, W. H., and Norman, D. A. Air Combat Maneuvering Performance Measurement State Space Analysis. Final Report, Contract F33615-80-C-0024. Brooks AFB, TX: HQ Air Force Human Resources Laboratory (AFSC), and Orlando, FL: Naval Training Equipment Center, July 1982.

Wooldridge, A. L. and Obermayer, R. W. Human-System Performance Diagnosis Methods for Automated and Adaptive Training. Contract N61339-80-D-0014. Thousand Oaks, CA: Vreuls Research Corporation, 1982.

APPENDIX A

PERFORMANCE MEASUREMENT OF AEROBATIC AND BASIC FIGHTER MANEUVERS

Aerobatic Maneuvers

Automated measurement of pilot/system performance during aerobatic maneuvers presents a complex challenge, but it is by no means an impossible task. During most such maneuvers, the desired position of the aircraft can be defined at all times according to such dimensions as heading, altitude, pitch, roll, and g-forces and by desired changes in these dimensions over time.

To simplify measurement, it is possible to divide each maneuver into discrete segments during which a defined sequence of aircraft states is expected. Deficiencies in pilot performance are seen as departures from the expected profile during each segment according to the set of measures being used.

Some of these segments are common to more than one maneuver and the measurement logic may, therefore, be transferred intact between maneuvers. Other segments are specific to a single maneuver.

It is also important to note that the measurement methodology may vary according to the characteristics of the airframe being used, the characteristics of the measurement system, and the reasons that the measurement is being done. Different airframes, for example, will have different recommended maneuver entry speeds and power settings. Differences between airframes, in some cases, will even dictate major differences in the way that a given maneuver is done. For example, the ability of the aircraft to sustain powered, inverted flight determines whether segments requiring zero or negative g-loading may be included in the maneuver technique.

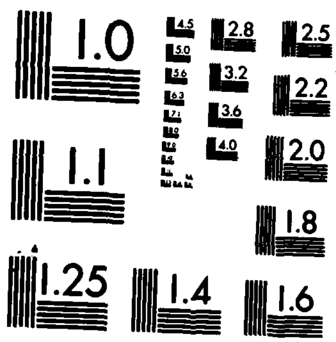
The characteristics of the measurement system may partially determine the measurement methodology. For example, there are several ways to determine, during a loop, when the aircraft nose passes the vertical position. The measurement capabilities of the system will be crucial in deciding whether this point is determined by a measure of aircraft pitch, heading, foreward velocity, or some other measure.

The reason that measurement is being taken may also impact the methodology to be used. One example of this is discussed under the section describing the Cuban eight.

The following sections describe how pilot performance might be measured during performance of a number of standard aerobatic maneuvers.

AD-A159 055 PERFORMANCE MEASUREMENT GUIDELINES FOR RESEARCH(U) 2/3
VREULS RESEARCH CORP THOUSAND OAKS CA D VREULS ET AL.
13 JUN 85 AFOSR-TR-85-0642 F49620-80-C-0058

UNCLASSIFIED F/G 5/9 NL


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3

THE LOOP

Probably the simplest aerobatic maneuver is the loop. A loop is essentially a 360 degree turn about the lateral axis of the aircraft in a vertical plane.

The proposed measurement methodology involves dividing the maneuver into six discrete segments, each having well defined start and stop logics. The initial segment begins when the instructor pilot signals the pilot to start the maneuver. The measurement system is initialized at this time. During the segment, the pilot is increasing or decreasing the aircraft speed to the prescribed maneuver entry speed for the specific aircraft and maneuver. Measurement of the segment ends when the aircraft has reached the entry speed. During this segment, aircraft speed and heading are the primary measurement variables.

The second segment begins when the aircraft has reached the maneuver entry speed and the pilot sets the throttle to the recommended entry power setting. The start logic, then, is of the form:

Airspeed = AS1 +/- Y knots AND Power = P1 +/- X units.

The segment ends when the aircraft pitch reaches vertical. This may be determined by a measurement of pitch angle or by the reversal of heading as:

Heading 2 = Heading 1 + 180 degrees +/- X degrees

depending on the parameters of the measurement system in use. During this segment, aircraft heading, pitch, airspeed, roll and power are of primary interest.

The third segment starts with the termination of segment 2 and terminates when pitch reaches 180 degrees (aircraft inverted). The variables of primary interest are the heading, pitch, roll and airspeed.

The fourth segment starts at the termination of segment 3 and ends when the pitch passes -90 degrees, either as determined by measurement of pitch angle or by the heading reversal. The measurement variable of primary importance are heading, pitch, roll and airspeed.

The fifth segment starts at the termination of segment 4 and ends with the return of the aircraft to level flight as determined by the pitch reaching 0 +/- X degrees. Variables of importance in this segment are heading, pitch, roll, airspeed and G's.

The sixth segment involves stabilization of the aircraft in level flight. it begins with the termination of segment five and ends when the instructor pilot terminates the exercise. Primary measurement variables are heading, roll, pitch, and airspeed.

THE IMMELMAN

The Immelman is essentially a 180 degree turn in a vertical plane. There are two accepted ways of performing the Immelman, depending on the capability of the aircraft to sustain inverted flight. The maneuver and measurement methodology described here are for an aircraft that is capable of inverted flight. The maneuver is divided into five segments.

Like the loop, segment 1 of the immelman consists of obtaining the recommended maneuver entry speed. It starts when the instructor pilot initializes the measurement system and starts the exercise. The segment terminates when the aircraft has reached the desired entry speed +/- X knots. Important variables are heading and airspeed.

The second segment of the immelman is also identical to the second segment of the loop. It begins when airspeed and power reach their recommended entry values and terminates when pitch reaches 90 degrees. Important variables are power, airspeed, roll, heading and pitch.

The third segment is similar, but not identical to that of the loop. It begins with the termination of segment two and ends when the aircraft pitch reaches the recommended value, slightly less than 180 degrees. Important variables for measurement are pitch, roll, heading and airspeed.

The fourth segment involves the roll from inverted to normal flight. It begins with the termination of segment three and ends when the aircraft has rolled upright to a bank angle of zero degrees +/- X degrees. Important variables are pitch, heading, roll angle and airspeed.

The fifth segment involves the stabilization of the aircraft in level flight and is, therefore, similar in many respects th the sixth segment of the loop. It begins with the termination of segment four and ends when the instructor pilot terminates the maneuver. Variables for measurement include pitch, heading, roll, airspeed and power.

THE CHANDELLE

The chandelle is a maximum performance climbing 180 degree turn during which the aircraft's kinetic energy is gradually reduced with recovery at near stall speed.

The maneuver is divided into four discrete measurement segments. Segment one is the maneuver entry. It begins when the instructor pilot triggers the measurement system to begin measurement of this maneuver and terminates when the airspeed reaches the recommended speed for this maneuver. This segment is identical to the first segment of the loop and immelman.

Segment two begins when the airspeed reaches the recommended entry airspeed +/- X knots AND the aircraft pitch angle is 0 degrees. It terminates when the heading has changed 90 degrees. During this segment, important measurement variables are heading, pitch, roll angle, sideslip, and airspeed.

The third segment begins with the termination of segment two and ends when the heading reaches 180 degrees beyond the entry heading or the bank angle reaches zero. Variables to be measured during this segment are heading, pitch, roll angle, sideslip and airspeed.

The fourth segment is the stabilization of the aircraft in level flight. It begins when the pitch reaches zero degrees +/- X degrees and it terminates when the instructor pilot terminates the maneuver. It is therefore essentially the same as the final measurement segment of the other maneuvers described in this section.

THE CUBAN EIGHT

The Cuban eight appears to be a very complex maneuver, however the performance measurement has much in common with the other maneuvers performed in the vertical plane, the loop and the immelman. In fact, six of the ten segments in the Cuban eight are common to the other maneuvers.

The first segment is used to bring the aircraft to the recommended maneuver entry speed. This segment is identical to those of the other maneuvers and is initiated by the instructor pilot. It terminates when the airspeed reaches the recommended entry speed +/- X knots.

The second segment is identical to that of the loop and the immelman. It begins when the power is at the recommended setting and the aircraft is at the recommended entry speed. It terminates when the pitch of the aircraft reaches 90 degrees (by whatever means this is determined.) Airspeed, power, pitch, roll and heading are important variables during this segment.

Measurement of the third segment is also identical to that for the loop and similar to that for the immelman. It begins with the termination of segment two and ends when the pitch reaches -180 degrees. Airspeed, pitch, roll and heading are important measurement variables during this segment.

The fourth segment begins with the end of the third segment and ends as the aircraft completes a 180 degree roll to return to a bank angle of zero degrees +/- X degrees. Important variables for measurement are pitch, roll and heading.

The diving segment 5 involves preparation for the pullup into the second loop of the Cuban eight. It begins with the end of segment 4 and ends as the pitch angle is increased through zero degrees. Pitch, roll, heading and airspeed are all important measurement variables.

After segment 5, the aircraft begins the second loop of the Cuban eight. This is a mirror image of the first loop and the measurement segments are identical except, of course, the aircraft headings are changed by 180 degrees.

The final segment, common to all of the maneuvers, involves the stabilization in normal flight. It begins after the second loop is completed

and the pitch angle reaches zero degrees +/- X degrees and ends when the instructor terminates the maneuver.

More detailed measurement of this maneuver may be needed, depending on the purpose of measurement. For example, it may be desired to measure the symmetry of the maneuver by measuring the altitudes at the top and bottom of each loop and the lengths of the various segments. This may provide a more difficult technical challenge for automated measurement.

THE LAZY EIGHT

The lazy eight maneuver is, essentially, a precision S-turn with carefully defined climbing and diving segments. Except for the entry and exit, the maneuver shares no common segments with any of the other maneuvers described in this section.

The first segment is initiated by the instructor pilot and terminates when the aircraft has reached the recommended maneuver entry speed +/- X knots. During this segment, measurement of heading, pitch and airspeed are important.

The second segment begins when the entry airspeed has been achieved and the pitch increases through zero degrees. The segment terminates when the heading has changed by 45 degrees. during this segment, the bank angle and pitch are constantly increasing. Important variables are pitch, bank, heading and airspeed, and various transforms of these across time.

The third segment starts with the end of the second and terminates when the heading has changed a total of 90 degrees. During this segment, the bank is constantly increasing but the pitch is constantly decreasing. Pitch should be zero degrees and bank at its steepest point at the termination of this segment. Important measurement variables are pitch, bank, heading and airspeed, and their transforms.

The fourth segment begins with the conclusion of the third and terminates when the heading has changed a total of 135 degrees. During this segment, the pitch and bank angle are constantly decreasing. Important variables, again, are pitch, bank, heading and airspeed, and their transforms.

The fifth segment begins with the end of the fourth segment and terminates when the heading has changed by a total of 180 degrees or when the bank angle reaches zero degrees. During the segment, bank is gradually decreasing and pitch is increasing toward zero. Ideally, at the conclusion of the segment, bank and pitch are both zero, heading has changed by 180 degrees and the airspeed is at the recommended entry speed. This segment completes the first half of the S-turn. Variables to be measured include pitch, bank, heading and airspeed.

The next four segments (6, 7, 8 and 9) are mirror images of the previous four as the aircraft completes a 180 degree turning maneuver in the opposite direction. Measurement is, therefore, identical.

During segment 10, the aircraft has completed the maneuver and becomes stabilized in normal flight.

Like the Cuban eight, it may be desired to increase the complexity of measurement in the lazy eight to test, for example, if the various parts of the maneuver are symmetrical in terms of altitude gain and loss, maximum bank and pitch angle and minimum and maximum airspeeds.

Basic Fighter Mananeuvers

Basic fighter maneuvers (BFM) are practiced during training to familiarize the pilot with the kinds of maneuvers required during air-to-air combat. The purpose of all of the BFM maneuvers is to improve one's position with respect to an adversary. These maneuvers can be divided into two main functional categories, offensive and defensive maneuvers.

The purpose of the offensive maneuvers is to bring the aircraft into the opponent's "cone of vulnerability," the position from which a weapon can successfully be launched. The purpose of defensive BFM maneuvers, on the other hand, is to move one's own cone of vulnerability away from the adversary.

Performance measurement of BFM must account for two different dimensions of performance: (1) the mechanical quality with which the pilot completes the maneuvers, and (2) the cognitive aspects of choosing the proper maneuver, timing it properly, and performing it according to the optimum parameters of speed, angle-of-attack, g-loading, heading and altitude change for the specific situation. These two dimensions should also be strongly related to the ultimate performance criterion: Did the maneuver achieve its purpose?

Most of the performance measurement techniques described in this text have involved identifying or developing a performance profile, describing the ideal path through space the aircraft should take and the control inputs required to keep the aircraft on that ideal profile. The performance of the pilot, according to a number of measured parameters, is then compared to that ideal profile in order to provide a score of performance.

During BFM, it is difficult to provide such ideal profiles. A Basic Fighter Maneuver (such as a high yo-yo) is not a single, strictly defined maneuver, but rather a large subset of similar, but not identical, maneuvers which serve the same general purpose but may be very different in terms of airspeed, altitude gain, g-loading, pitch rate, roll rate and timing of performance. The optimum values of these parameters are determined by the relative positions of the two aircraft and the intentions of the pilot.

Because BFM involves no "standardized" maneuvers and the parameters of a maneuver depend on the positional relationship between two aircraft,

there can be no firmly established maneuver profiles. For this reason, meaningful automated measurement of the pilot's mechanical aircraft handling performance according to an assumed optimal set of parameters is difficult and does not account for large amounts of performance variance.

In fact, one measurement model involving specific mechanical control measures which proved rather effective for measurement of performance during free air combat engagements had virtually no utility for measurement during BFM maneuvers (Kelly, Wooldridge, Hennessy, Vreuls, Barnebey, Cotton and Reed, 1979.) Performance measurement of specific aircraft parameters during BFM, therefore, should concentrate on a tolerance band approach simply to assure that the important parameters such as g-loading, angle of attack and airspeed remain within accepted bounds during the maneuver.

This conclusion may not apply, however, to future "intelligent" performance measurement systems which might model an ideal pilot's thought process and, for a given situation of aircraft position and energy, determine the ideal profile and then contrast the pilot's performance to that ideal. Such an artificial intelligence capability does not yet exist, though numerous "iron pilot" programs are in operation. None of these claim to provide an accurate model of an ideal pilot's tactical planning and maneuvering.

Measurement methods which concentrate on the outcomes, rather than techniques, of BFM maneuvers have been more effective (Moore, Madison, Sepp, Stracener and Coward, 1979; Simpson and Oberle, 1977; Wooldridge, Kelly, Obermayer, Vreuls, Nelson and Norman, 1982.) Measurements of the quickness and precision with which the pilot can improve his situation in terms of relative aircraft position measures probably provide the most effective single approach to performance measurement of BFM maneuvers.

The relevant measures upon which to base such a methodology are the aspect angle, line of sight angle, range, and closure rate between aircraft, as well as the energy states of the aircraft.

The aspect angle for Aircraft A is defined as the angular distance between the longitudinal axis of Aircraft B and the line of sight between Aircraft A and Aircraft B. It can be any number between 0 degrees and 180 degrees. With zero degree aspect angle, the pilot is looking directly up his adversary's tailpipe and with 180 degrees aspect he is looking at his adversary's nose.

The line of sight angle for Aircraft A is defined as the angular distance between the longitudinal axis of Aircraft A and the line of sight between aircraft A and Aircraft B. This can be any number between 0 degrees and 180 degrees. With a zero degree line of sight angle, the pilots adversary is directly ahead at the 12 o'clock position and with a 180 degree line of sight, the adversary is directly behind at the six o'clock position.

Range is simply the linear distance between Aircraft A and Aircraft B. Closure rate is the rate of change (e.g., feet per second) in the range.

Energy level is the total kinetic and potential energy of an aircraft due to the combination of airspeed and altitude.

Calculation of the aspect angle, line of sight, range, closure, and energy states of the two aircraft, of course, requires the measurement and transformation of more basic position data for each aircraft. This includes the instantaneous position of each aircraft in latitude, longitude and altitude as well as the heading, roll and pitch angles.

The following sections describe a sample of BFM maneuvers and some candidate measurement methods for them:

HARD TURN

The hard turn is used to quickly rotate one's cone of vulnerability away from an attacker who has achieved a position of advantage and is an immediate threat. The intensity of the turn is determined by the position of the adversary; the tightness of the turn may be changed after the turn is initiated in response to the opponent's moves. The entry to the hard turn is usually at a prebriefed g-load or angle-of-attack which will allow continuation of the turn without a loss of energy.

The outcome of the maneuver can be measured by comparing aspect angle, line of sight angle, range, closure and energy measures before and after the maneuver. In addition, the angle-of-attack or g-loading should be tracked throughout the maneuver to determine whether the parameters remained within accepted bounds. In addition, sideslip should be tracked during all high angle-of-attack maneuvers.

BREAK TURN

The break turn is a maximum rate turn which employs the highest available g-loading. The turn quickly dissipates energy and is, therefore, most applicable to last ditch defensive maneuvering. During training, a prebriefed g-loading is used for the break turn.

The effectiveness of the maneuver, again, is determined by measuring the aspect angle, line of sight angle, range, closure, and energy before and after the maneuver. G-loading or angle-of-attack and sideslip should also be tracked to assure that they remained within desired tolerances during the maneuver.

CORNER TURN

The corner turn (a special case of the break turn) is the tightest turn that can be made by a specific aircraft type at a given altitude and weight without stalling or sustaining aircraft structural damage. The corner turn is only possible at one specific aircraft speed (the corner velocity) at the given altitude. The corner turn may be used during both offensive and defensive maneuvering.

Because it is flown according to defined parameters, the corner turn is slightly different than the other BFM maneuvers described here and more

mechanical measurement may be used. In addition to the position measures (aspect angle, line of sight angle, range, closure and energy) important measures include altitude, airspeed, g-loading and sideslip.

REVERSAL

A reversal is a rapid change in the direction of a turn, made by a defender in an attempt to exploit an attacker's overshoot on the outside of the defender's turn. It is accomplished by a rapid roll, loaded or unloaded depending on the attackers position, into the opposite direction of the original turn.

Performance measurement during the reversal should concentrate primarily on the outcome of the maneuver in terms of changes in aspect angle, line of sight angle, range, closure and energy. Depending on the purpose of measurement, however, other measures such as roll rate, g's, power and pitch during the reversal maneuver may be desirable.

HIGH YO-YO

The high yo-yo is performed by the attacking aircraft in order to prevent overtaking the defender during a turn. In this maneuver, the attacking aircraft pulls upward, out of the plane of the turn, and then back down into the turning plane. This allows the attacker to dissipate energy and to travel a slightly longer flightpath than the defender, thereby preventing an overshoot.

Performance measurement during the high yo-yo is primarily based on the outcome of the maneuver and whether it accomplished its desired goals of decreasing aspect angle and line of sight angle or maintain a low aspect and line of sight while greatly decreasing the closure rate. Thus, the most important performance indicator is the difference between the aspect, line of sight, and closure rates at the inception and termination of the maneuver. Other measures of some importance are airspeed, altitude and g-loading during the maneuver.

LOW YO-YO

The low yo-yo is an acceleration and cutoff maneuver, performed to allow the attacking aircraft to close on a maneuvering defender. In this maneuver, the attacking aircraft pushes downward out of the plane of the defenders maneuver and, while accelerating, turns to cut across the tail of the defender. The attacker then pulls upward to intercept the defender.

Performance measurement of the low yo-yo is very similar to that for the high yo-yo since both maneuvers are involved with optimizing the closure rate while maintaining or decreasing aspect and line of sight angles. Aspect angle, line of sight angle, range, and closure should be measured before and after the maneuver with the quality of the maneuver indicated by the differences. Other important variables may be airspeed, altitude and g-loading.

SCISSORS

The scissors is a slow speed, high angle-of-attack series of turn reversals by a pair of adversaries, neither of whom has a positional advantage. It generally begins when an attacking aircraft overshoots, the defender reverses, and the pilots find themselves canopy to canopy. They each continue reversals, waiting for the other to move in front or separate.

Performance measurement of the scissors is especially complex because the entry is usually unintentional and there is no well defined termination point for the maneuver. Theoretically, with evenly matched pilots and aircraft, the scissors could continue until one or both aircraft ran out of fuel. Once the pilot recognizes that he has entered a scissors, he may elect either to remain in the maneuver, maintaining low speeds and high angles of attack, until he is able to decrease his aspect and line of sight, or he may choose to attempt to separate.

During the scissors, measures of aspect angle and line of sight angle are of utmost importance. Other important variables may be airspeed, angle of attack, power, pitch and roll rate. If the pilot elects to separate from the maneuver, the aspect and line of sight angles at the point of separation are also important.

A variation of the scissors is the vertical rolling scissors in which both aircraft find themselves in a tight descending spiral. This situation is considered more of a predicament than a bona fide BFM maneuver, and is not entered intentionally.

BARREL ROLL ATTACK

The barrel roll attack is similar to the high yo-yo in that it allows the attacker to cut off and intercept a defender while controlling for a possible overshoot. The maneuver begins at relatively long range and with the attacker at a high airspeed and involves a vertical roll by the attacker to control the potential overshoot while conserving energy. The aircraft exits the vertical roll in an improved position of aspect angle, line of sight angle and range.

Again, the primary measurement criterion is the outcome of the maneuver in terms of aspect angle, line of sight angle, range, closure, and comparative energy states, between the beginning and end of the maneuver. Also important is measurement of G's, pitch, roll rate, airspeed and altitude.

1-V-1 FREE ENGAGEMENTS

During training, BFM is practiced primarily against a target aircraft that is flying a predictable profile. This allows the student pilot to become familiar with the required maneuvers without having to be concerned about the adversary's combat tactics. The student is allowed to fly a complete maneuver and to recognize the maneuver outcomes in terms of any improvement in relative position.

During 1-v-1 free engagements, the student takes one step closer to the real world air combat situation. During these free engagements, each aircraft is maneuvering in response to the other and both are trying to obtain and maintain a position of advantage. While air combat engagements of this nature would be rare under current air warfare doctrine, this phase simulates the classic one on one dogfights of previous eras.

While BFM forms the foundation for maneuvering during the free engagements it is rare that a textbook BFM maneuver is flown from start to completion. Pilot performance during free engagements involves smooth transitions between portions of BFM maneuvers as a constantly changing stream of responses to the adversary.

Measurement of performance during free engagements, therefore, is one step farther removed from classical, profile-based measurement than the other BFM maneuvers. Performance during 1-v-1 free engagements has a number of dimensions that must be captured empirically to provide a complete performance metric. During interviews with air combat experts, Kelly, et al.(1979) developed a taxonomy of skills, traits and performance dimensions important to pilot performance. These included:

Personal traits

- Aggressiveness
- Decisiveness
- Situation awareness

Knowledge and skills

- Knowledge of weapons and tactics
- Knowledge of and ability to apply BFM
- Basic flying skill

Performance indicators

- Maintains offensive position
- Wins engagements

The study then completed a detailed performance analysis of a large number of 1-v-1 free engagements and, through multivariate statistical techniques, found that several types of measures, averaged or tallyed across an entire engagement, could be used as indicators of performance. These included measures of energy management, control activity (especially throttle and speedbrake,) relative aircraft positions, aircraft maneuvering activity, and weapons success.

Measurement of relative aircraft positions appears to be the most crucial factor in the empirical measurement of 1-v-1 free engagements both as a means of measuring performance, itself, and as a means of segmenting an engagement for the application of other performance measures.

Some relative position measures that have been found to be important in themselves include:

Offensive time. The percentage of an engagement that a pilot was in an offensive position as defined by aspect angle, line of sight angle, and range.

Defensive time. The percentage of time during an engagement that a pilot was in a defensive position as defined by aspect angle, line of sight, and range.

Offensive time/Defensive time. The ratio of offensive time to defensive time.

Gun tracking error. The average mil error while in a gun tracking position.

Time opponent in sight. Based on the field of view from the cockpit, the percentage of time during an engagement that an opponent is visible to the pilot.

Time in a weapons envelope. The percentage of time during an engagement that the pilot is in his adversary's cone of vulnerability with a line of sight angle of near zero.

Aspect angle + Line of sight angle. The sum of the aspect angle and line of sight angle provides a parsimonious indicator of the position of relative advantage or disadvantage at a given instant. As the sum approaches zero degrees, advantage is greater and as it approaches 360 degrees, disadvantage is greater.

Relative aircraft position may also serve as a means of segmenting a 1-v-1 engagement into subparts for more precise measurement. This approach has been referred to (Wooldridge, et al., 1982) as a TACSPACE model. Combinations of aircraft aspect angles, line of sight angles, ranges, and elevations above or below the fighter can be represented with a three dimensional cube. Based on the measurement data to be analyzed, this structure can be modified to combine cells as needed according to the specific measurement situation.

This approach assumes that at a given combination of aspect, line of sight, range, and elevation, the viable options available to a pilot in terms of maneuvering and tactics will be greatly limited. Therefore, within a given TACSPACE cell, variance within a given control measure due to aircraft position effects should be greatly reduced.

REFERENCES

Kelly, M. J., Wooldridge, L., Hennessy, R. T., Vreuls, D., Barneby, S. F., Cotton, J. C., and Reed, J. C. Air Combat Maneuvering Performance Measurement. NAVTRAEQUIPCENIH-35/AFHRL-TR-79-3. Orlando, FL: Naval Training Equipment Center, and Brooks AFB, TX: Air Force Human Resources Laboratory, September, 1979.

Moore, S. B., Madison, W. G., Sepp, G. D., Stracener, J. T. and Coward, R. E. Air Combat Training: Good Stick Index Validation. AFHRL-TR-79-15. Williams AFB, AZ: Flying Training Division, Air Force Human Resources Laboratory, May 1979.

Simpson, W. R. and Oberle, R. A. The Numerical Analysis of Air Combat Engagements Dominated by Maneuvering Performance. TM 77-2 SA. Patuxent River, MD: Naval Air Test Center, June 1977.

Wooldridge, L., Kelly, M. J., Obermayer, R. W., Vreuls, D., Nelson, W. H., and Norman, D. A. Air Combat Maneuvering Performance Measurement State Space Analysis. Final Report, Contract F33615-80-C-0024. Brooks AFB, TX: HQ Air Force Human Resources Laboratory (AFSC), and Orlando, FL: Naval Training Equipment Center, July 1982.

APPENDIX B

FORTRAN TRANSFORMATION PROGRAMS

This appendix consists of a series of FORTRAN programs discussed in Section 6.

THDAT.FOR
LSTRP3.FOR
STRIP.FOR
THMEAS.FOR
AMPDIST.FOR
CUMDIST.FOR
MISR.FOR
PFOURIER.FOR
MULTR.FOR
CROSS.FOR
PLOT.FOR
RANDU.FOR

```
C -------------------------
C     PROGRAM THDAT
C -------------------------
C
C PROGRAM TO GENERATE SAMPLE
C TIME HISTORY DATA
C
 REAL I
 TUPI = 2. * 3.141592654
 IF (IOWRIT(5,0,0,"THDAT.FIL")) STOP
 DO 500 I=0.0,60.2,0.2
C
C GENERATE TARGET DATA
C
 TGT = 2.*SIN (TUPI*0.5*I) + 2.5*SIN(TUPI*0.4*I)
     1 +3.33333333*SIN(TUPI*0.3*I) + 5.*SIN(TUPI*.2*I)
     2 + 10.*SIN(TUPI*0.1*I)
C
C GENERATE HUMAN RESPONSE DATA
C (FIRST TERM IS HIGH FREQU. NOISE, AND
C HIGHEST TWO FREQUENCIES SHOW SOME PHASE LAG)
C
C INSERT ONE STEP OF DELAY INTO OPERATOR RESPONSE
C
 I=I-0.2
C
 HR = 0.5*SIN(TUPI*0.75*I - 1.0)
     1 + 1.5*SIN(TUPI*0.5*I -1.0) +2.2*SIN(TUPI*.4*I -.75)
     2 +3.*SIN(TUPI*0.3*I) +4.9*SIN(TUPI*0.2*I)
     3 + 9.8*SIN(TUPI*0.1*I)
C
C RESET VALUE OF I
C
 I=I+0.2
C
C
C COMPUTE ERROR
C
 ERR = TGT - HR
C
C OUTPUT DATA
C
 WRITE (1,100) I, TGT, HR, ERR
 WRITE (5,100) I, TGT, HR, ERR
100 FORMAT (F5.1,8F8.3)
C
C END OF LOOP
C
500  CONTINUE
 IF(IOCLOS(5)) STOP
 STOP
 END
C ----------------------------
C     PROGRAM LSTRP3
C ----------------------------
```

```
C READ TIME HISTORY DATA
C AND PLOT IN STRIP CHART FORMAT
 DIMENSION X(151),Y(151),Z(151)
 INTEGER*2 I,N
C
C OPEN THE FILE
 IF (IOREAD (5,0,0,"THDAT.FIL")) GOTO 999
C READ DATA
 DO 50 I=1,301
 READ (5,100,ENDFILE=50) T,TGT,HR,ERR
 X(I)=TGT;Y(I)=HR;Z(I)=ERR
50 CONTINUE
 IF(IOCLOS(5)) STOP 3
 CALL LSTRP3(X,Y,Z,151,-25.,25.,0.,30.)
60 STOP
999 WRITE (1) "ERROR"
   STOP
100 FORMAT (F5.1,3F8.3)
 END
C
C
C
 SUBROUTINE LSTRP3(V,W,X,N,XMIN,XMAX,TMIN,TMAX)
C
C SUBROUTINE PRINTS A STRIP CHART FOR
C THREE VARIABLE WITH PLOTTING SCALE SET
C FOR COMPUTER SIZE PAPER OR 16.7 CHAR/IN.
C
C TVALUES MUST BE AT REGULAR INTERVALS, E.G.,
C TIME
C X= ARRAY TO BE PLOTTED IN STRIP CHART FORM
C ALSO VARIABLES V AND W.
C N = NUMBER OF POINTS TO BE PLOTTED
C XMIN, XMAX = RANGE OF X-VALUES
C TIMIN, TMAX = RANGE OF PARAMETER AGAINST WHICH
C X IS TO BE PLOTTED AT REGULAR INTERVALS
C (LONG DIMENSION OF THE PAPER)
C
 DIMENSION X(151),IA(101),XLAB(21),YLAB(61)
 DIMENSION V(151), W(151)
 INTEGER*2 N,I,J,J1,I1,I2,IA
 DATA ISP,IPR,IS/1H ,1H*,1HX/
 DATA IPL/1H+/
20 FORMAT (10X,1H+,101A1)
30 FORMAT (1X,F8.2,1X,1H¼,101A1)
40 FORMAT (10X,2H+I,10(10H+++++++++I))
50 FORMAT (4X,11(3X,F7.2))
C
C COMPUTE SPACING
C
 XSP = 100.
 YSP = N-1
 DX = XSP/(XMAX-XMIN)
 DY = YSP/(TMAX-TMIN)
C
```

```
C COMPUTE REQUIRED LABELS
C
 DO 400 I=1,101,5
 J = I/5 + 1
400 XLAB(J) = (I-1)/DX+XMIN
 DO 410 I=1,N,5
 J=I/5+1
410 YLAB(J) = (I-1)/DY+TMIN
C
C WRITE BEGINNING X LABELS
C
 WRITE (4,50) (XLAB(I),I=1,21,2)
 WRITE (4,40)
C
C PLOT N VALUES OF X
C
 DO 500 J1=1,N
 IF(W(J1).GT.XMAX) W(J1)=XMAX
 IF(W(J1).LT.XMIN) W(J1)=XMIN
 IF(V(J1).GT.XMAX) V(J1)=XMAX
 IF(V(J1).LT.XMIN) V(J1)=XMIN
 IF(X(J1).GT.XMAX) X(J1)=XMAX
 IF(X(J1).LT.XMIN) X(J1)=XMIN
C
C INITIALIZE PLOT ARRAY WITH SPACES
C
 DO 200 I=1,101
200 IA(I)=ISP
 XX=DX*(X(J1)-XMIN)+1.5
 I=IFIX(XX)
C
C CHANGE PROPER SPACE TO A PLOT SYMBOL
C
 IA(I) = IPR
 I=DX*(V(J1)-XMIN)+1.5
 IA(I)=IS
 I=DX*(W(J1)-XMIN)+1.5
 IA(I)=IPL
C
C WRITE A LINE OF PLOT INFORMATION
C
 DO 450 I1=1,100
 I2 = 102-I1
 IF (IA(I2).NE.ISP) GOTO 475
450 CONTINUE
475 IF (MOD((J1-1),5).EQ.0) GOTO 490
 WRITE (4,20) (IA(I),I=1,I2)
 GOTO 500
490 WRITE (4,30) YLAB (J1/5+1),(IA(I),I=1,I2)
500 CONTINUE
C
C WRITE ENDING X LABELS
C
 WRITE (4,40)
 WRITE (4,50) (XLAB(I),I=1,21,2)
 END
```

```
C TEST SUBROUTINE STRIP
 DIMENSION X(301)
 INTEGER*2 I,N
C
C OPEN THE FILE
 IF (IOREAD (5,0,0,"THDAT.FIL")) GOTO 999
C READ DATA
 DO 50 I=1,301
 READ (5,100,ENDFILE=50) T,TGT,HR,ERR
50 X(I)=TGT
 CALL STRIP (X,301,-25.,+25.,0.,60.)
   IF (IOCLOS(5)) GOTO 999
60 STOP
999 WRITE (1) "ERROR"
   STOP
100 FORMAT (F5.1,3F8.3)
 END
C
C
C
 SUBROUTINE STRIP(X,N,XMIN,XMAX,TMIN,TMAX)
C TVALUES MUST BE AT REGULAR INTERVALS, E.G.,
C TIME
C X= ARRAY TO BE PLOTTED IN STRIP CHART FORM
C N = NUMBER OF POINTS TO BE PLOTTED
C XMIN, XMAX = RANGE OF X-VALUES
C TIMIN, TMAX = RANGE OF PARAMETER AGAINST WHICH
C X IS TO BE PLOTTED AT REGULAR INTERVALS
C (LONG DIMENSION OF THE PAPER)
C
 DIMENSION X(301),IA(51),XLAB(11),YLAB(61)
 INTEGER*2 N,I,J,J1,I1,I2,IA
 DATA ISP,IPR,IS/1H ,1H*,1HX/
20 FORMAT (10X,1H+,80A1)
30 FORMAT (1X,F9.2,1H¼,80A1)
40 FORMAT (10X,2H+I,5(10H+++++++++I))
50 FORMAT (6X,6F10.2)
C
C COMPUTE SPACING
C
 XSP = 50.
 YSP = N-1
 DX = XSP/(XMAX-XMIN)
 DY = YSP/(TMAX-TMIN)
C
C COMPUTE REQUIRED LABELS
C
 DO 400 I=1,51,5
 J = I/5 + 1
400 XLAB(J) = (I-1)/DX+XMIN
 DO 410 I=1,N,5
 J=I/5+1
410 YLAB(J) = (I-1)/DY+TMIN
C
C WRITE BEGINNING X LABELS
```

```
C
 WRITE (4,50) (XLAB(I),I=1,11,2)
 WRITE (4,40)
C
C PLOT N VALUES OF X
C
 DO 500 J1=1,N
C
C INITIALIZE PLOT ARRAY WITH SPACES
C
 DO 200 I=1,51
200 IA(I)=ISP
 XX=DX*(X(J1)-XMIN)+1.5
 I=IFIX(XX)
C
C CHANGE PROPER SPACE TO A PLOT SYMBOL
C
 IA(I) = IPR
C
C WRITE A LINE OF PLOT INFORMATION
C
 DO 450 I1=1,50
 I2 = 52-I1
 IF (IA(I2).NE.ISP) GOTO 475
450 CONTINUE
475 IF (MOD((J1-1),5).EQ.0) GOTO 490
 WRITE (4,20) (IA(I),I=1,I2)
 GOTO 500
490 WRITE (4;30) YLAB (J1/5+1),(IA(I),I=1,I2)
500 CONTINUE
C
C WRITE ENDING X LABELS
C
 WRITE (4,40)
 WRITE (4,50) (XLAB(I),I=1,11,2)
 END
C ----------------------------------------
C     PROGRAM THMEAS
C ----------------------------------------
C
C READ THE THDAT.FIL FILE AND COMPUTE
C TIME-HISTORY MEASUREMENTS
C  TIME ON TARGET
C  ZERO CROSSINGS
C  PEAK VALUE
C  AVERAGE ERROR
C  ABSOLUTE AVERAGE ERROR
C  MEAN SQUARE ERROR
C  ROOT MEAN SQUARE ERRORS
C  REVERSALS
C
 REAL MS
C
C INITIALIZE
C
```

```
 TOT = 0.0
 BAND = 1.0
 ZCROSS = 0.0
 PEAK = 0.0
 COUNT = 0.0
 AE = 0.0
 AAE = 0.0
 MS = 0.0
 REVERS = 0.0
 RTOL = 0.2
 TEND = 60.0
C
C OPEN THE FILE
C
 IF(IOREAD(5,0,0,"THDAT.FIL"))GOTO 999
C
C
C READ DATA
1 CONTINUE
C
 READ (5,100,ENDFILE=50) T,TGT,HR,ERR
100 FORMAT (F5.1,3F8.3)
C
C COMPUTE TIME ON TARGET
C
 IF (ABS(ERR) .LE. BAND) TOT=TOT+1.0
C
C
C COMPUTE ZERO CROSSINGS
C
 IF(T) 10,10,11
10 ALAST = ERR/ABS(ERR)
 GOTO 12
11 A1 = ERR/ABS(ERR)
 IF(ABS(A1-ALAST)-.2) 14,14,13
13 ZCROSS = ZCROSS + 1.0
14 ALAST = A1
12 CONTINUE
C
C COMPUTE PEAK VALUE
C
 IF(ABS(ERR) .GT. ABS(PEAK)) PEAK=ERR
C
C COMPUTE AVERAGE ERROR
C
 AE = AE + ERR
 COUNT = COUNT + 1.0
C
C COMPUTE ABSOLUTE AVERAGE ERROR
C
 AAE = AAE + ABS(ERR)
C
C COMPUTE MEAN SQUARE ERROR
C
 MS = MS + ERR*ERR
```

```
C
C
C COMPUTE REVERSAL COUNT
C
C PICK UP FIRST VALUES FOR DERIVATIVE CALC.
 IF (T) 21,21,22
21 TLAST = TGT
 HLAST = HR
 GOTO 26
C ALSO SKIP OVER LAST VALUE
22 IF (T - TEND) 23,26,26
C COMPUTE SIGN OF DERIVATIVES
23 TDOT = TGT -TLAST
 STDOT = TDOT/ABS(TDOT)
 HDOT = HR - HLAST
 SHDOT = HDOT/ABS(HDOT)
C COUNT A REVERSAL IF DERIVATIVES ARE NOT
C THE SAME SIGN
 IF(ABS(STDOT - SHDOT) - RTOL) 25,25,24
24 REVERS = REVERS + 1.0
C UPDATE "LAST VALUE" FOR DERIVATIVE CALC
25 HLAST = HR
 TLAST = TGT
26 CONTINUE
C
 GOTO 1
50 CONTINUE
C
C FINISH CALCULATIONS
C
 AE = AE/COUNT
 AAE = AAE/COUNT
 COUNT = COUNT - 1.0
 MS = MS/COUNT
 RMS = SQRT(MS)
C
C OUTPUT RESULTS
C
 WRITE (4,110) TOT
110 FORMAT ( " TIME (SAMPLES) ON TARGET = ", F8.2)
 WRITE (4,111) ZCROSS
111 FORMAT ( " ZERO CROSSINGS = ", F8.0 )
 WRITE (4,112) AE
112 FORMAT (" AVERAGE ERROR = ", F8.2 )
 WRITE (4,113) AAE
113 FORMAT (" ABSOLUTE AVERAGE ERROR = ", F8.2)
 WRITE (4,114) MS
114 FORMAT (" MEAN SQUARE ERROR = ", F8.2)
 WRITE (4,115) RMS
115 FORMAT (" ROOT MEAN SQUARE ERROR = ", F8.2)
 WRITE (4,116) PEAK
116 FORMAT (" PEAK VALUE = ", F8.2)
 WRITE (4,117) REVERS
117 FORMAT (" NUMBER OF REVERSALS = ", F8.2)
C
```

```
C CLOSE FILE
C
 IF (IOCLOS(5)) GOTO 999
60 STOP
999 WRITE (1) "ERROR"
 END
C ---------------------------------------
C      PROGRAM AMPDIST
C ---------------------------------------
C
C COMPUTE AN DISTRIBUTION OF
C AMPLITUDES AND PRODUCE A PLOT
C
 DIMENSION A(112), B(112)
C
C THE RANGE OF THE DISTRIBUTION IS BETWEEN
C -BMAX AND +BMAX
C
 BMAX = 10.0
 BMIN = -10.0
C
C INITIALIZE ARRAYS TO BE PLOTTED
 DO 1 K=1,112
 AK=K-1
 A(K)=0.0
1 B(K)=BMAX* (-1.0 + 2.0*AK/100.)
C
 EN = 0.0
C
C OPEN THE FILE
 IF (IOREAD (5,2,0,"THDAT.FIL")) STOP 1
 IF (IOWRIT(6,2,0,"AMPDIST.FIL")) STOP 8
C READ DATA
2 CONTINUE
 READ (5,100,ENDFILE=50) T,TGT,HR,ERR
 EN = EN + 1.0
100 FORMAT (F5.1,3F8.3)
C
C COMPUTE CUMUL. DISTRIBUTION
C CHANGE DENSITY OF PLOT BY STEP SIZE
C ON THE FOLLOWING DO-LOOP
C
C SET STEP = 10 FOR THIS EXAMPLE (SET BY
C TRIAL AND ERROR
 DO 3 I= 1,101,10
C NOTE B(I+STEP) AND A(I+HALF-STEP) IN THE
C NEXT INSTRUCTION
 IF((B(I).LE.ERR).AND.(B(I+10).GT.ERR))A(I+5)=A(I+5)+1.0
3 CONTINUE
 GOTO 2
50 CONTINUE
C
C PLOT
C
 CALL PLT (B,A,101,BMIN,BMAX,0.,100.)
```

```
C
C CLOSE FILE
C
      IF (IOCLOS(5)) STOP 2
      IF(IOCLOS(6)) STOP 9
60    STOP 3
      END
C
C
C
      SUBROUTINE PLT (X,Y,N,XMIN,XMAX,YMIN,YMAX)
      DIMENSION X(112),Y(112),IA(51,51),XLAB(11),YLAB(11)
      DATA ISP,IPR,IS/1H ,1H* ,1HX/
20    FORMAT (10X,1H+,80A1)
30    FORMAT (1X,E9.2,1H-,80A1)
40    FORMAT (10X,2H+I,5(10H+++++++++I))
50    FORMAT (6X,6E10.2)
      XSP =50.
      YSP=XSP
      DX=XSP/(XMAX-XMIN)
      DY=YSP/(YMAX-YMIN)
      DO 200 I=1,51
      DO 200 J=1,51
200   IA(I,J)=ISP
      DO 300 K=1,N
      I=DX*(X(K)-XMIN)+1.5
      J=DY*(Y(K)-YMIN)+1.5
      IA(I,J)=IPR
300   CONTINUE
      DO 400 I=1,51,5
      J=I/5+1
      XLAB(J)=(I-1)/DX+XMIN
      XLAB(1)=XLAB(1)
      YLAB(J)=(I-1)/DY+YMIN
400   CONTINUE
      DO 500 J1=1,51
      J=52-J1
      DO 450 I1=1,50
      I2=52-I1
      IF(IA(I2,J).NE.ISP) GO TO 475
450   CONTINUE
475   IF(MOD((J-1),5).EQ.0) GO TO 490
      WRITE (6,20) (IA(I,J),I=1,I2)
      GO TO 499
490   WRITE (6,30) YLAB(J/5+1),(IA(I,J),I=1,I2)
499   CONTINUE
500   CONTINUE
      WRITE (6,40)
      WRITE (6,50) (XLAB(I),I=1,11,2)
      RETURN
      END
C --------------------------------------
C      PROGRAM CUMDIST
C --------------------------------------
C
```

```
C COMPUTE AN CUMULATIVE DISTRIBUTION OF
C AMPLITUDES AND PRODUCE A PLOT
C
 DIMENSION A(101), B(101)
C
C INITIALIZE ARRAYS TO BE PLOTTED
C THE RANGE OF THE DISTRIBUTION IS BETWEEN
C -BMAX AND +BMAX.
 BMAX = 10.0
 DO 1 K=1,101
 AK=K-1
 A(K)=0.0
1 B(K) = BMAX * (-1.0 + 2.*AK/100.)
 EN = 0.0
C
C OPEN THE FILE
 IF (IOREAD (5,2,0,"THDAT.FIL")) STOP 1
 IF (IOWRIT (6,2,0,"CUMDIST.FIL")) STOP 9
C READ DATA
2 CONTINUE
 READ (5,100,ENDFILE=50) T,TGT,HR,ERR
 EN = EN + 1.0
 WRITE (1,100) T
100 FORMAT (F5.1,3F8.3)
C
C COMPUTE ACCUM. DISTRIBUTION
C
 DO 3 I= 1,101
 IF (ERR .LT. B(I) ) A(I)=A(I) + 1.0
3 CONTINUE
 GOTO 2
50 CONTINUE
C
C CONVERT SO THAT ABSCISSA CAN BE READ
C IN PERCENT
C
 DO 4 L= 1,101
 AL=L
 WRITE (1,100) AL,B(L),A(L)
4 A(L)= 100.0* A(L)/EN
C
C PLOT
C
 BMIN=(-1.)*BMAX
 CALL PLT (B,A,101,BMIN,BMAX,0.,100.)
C
C CLOSE FILE
C
 IF (IOCLOS(5)) STOP 2
 IF (IOCLOS(6)) STOP 8
60 STOP 3
 END
C
C
C
```

```
C
C     ...............................................................
C
C        SUBROUTINE MISR
C
C        PURPOSE
C           COMPUTE MEANS, STANDARD DEVIATIONS, SKEWNESS AND KURTOSIS,
C           CORRELATION COEFFICIENTS, REGRESSION COEFFICIENTS, AND
C           STANDARD ERRORS OF REGRESSION COEFFICIENTS WHEN THERE ARE
C           MISSING DATA POINTS.  THE USER IDENTIFIES THE MISSING DATA
C           BY MEANS OF A NUMERIC CODE.  THOSE VALUES HAVING THIS CODE
C           ARE SKIPPED IN COMPUTING THE STATISTICS.  IN THE CASE OF THE
C           CORRELATION COEFFICIENTS, ANY PAIR OF VALUES ARE SKIPPED IF
C           EITHER ONE OF THEM ARE MISSING.
C
C        USAGE
C           CALL MISR (NO,M,X,CODE,XBAR,STD,SKEW,CURT,R,N,A,B,S,IER)
C
C        DESCRIPTION OF PARAMETERS
C           NO   - NUMBER OF OBSERVATIONS
C           M    - NUMBER OF VARIABLES
C           X    - INPUT DATA MATRIX OF SIZE NO X M.
C           CODE - INPUT VECTOR OF LENGTH M, WHICH CONTAINS A NUMERIC
C                  MISSING DATA CODE FOR EACH VARIABLE. ANY OBSERVATION
C                  FOR A GIVEN VARIABLE HAVING A VALUE EQUAL TO THE CODE
C                  WILL BE DROPPED FOR THE COMPUTATIONS.
C           XBAR - OUTPUT VECTOR OF LENGTH M CONTAINING MEANS
C           STD  - OUTPUT VECTOR OF LENGTH M CONTAINING STANDARD DEVI-
C                  ATIONS
C           SKEW - OUTPUT VECTOR OF LENGTH M CONTAINING SKEWNESS
C           CURT - OUTPUT VECTOR OF LENGTH M CONTAINING KURTOSIS
C           R    - OUTPUT MATRIX OF PRODUCT-MOMENT CORRELATION
C                  COEFFICIENTS.  THIS WILL BE THE UPPER TRIANGULAR
C                  MATRIX ONLY, SINCE THE M X M MATRIX OF COEFFICIENTS
C                  IS SYMMETRIC. (STORAGE MODE 1)
C           N    - OUTPUT MATRIX OF NUMBER OF PAIRS OF OBSERVATIONS USED
C                  IN COMPUTING THE CORRELATION COEFFICIENTS.  ONLY THE
C                  UPPER TRIANGULAR PORTION OF THE MATRIX IS GIVEN.
C                  (STORAGE MODE 1)
C           A    - OUTPUT MATRIX (M BY M)  CONTAINING INTERCEPTS OF
C                  REGRESSION LINES (A) OF THE FORM Y=A+BX.  THE FIRST
C                  SUBSCRIPT OF THIS MATRIX REFERS TO THE INDEPENDENT
C                  VARIABLE AND THE SECOND TO THE DEPENDENT VARIABLE.
C                  FOR EXAMPLE, A(1,3) CONTAINS THE INTERCEPT OF THE
C                  REGRESSION LINE FOR TWO VARIABLES WHERE VARIABLE 1
C                  IS INDEPENDENT AND VARIABLE 3 IS DEPENDENT.  NOTE
C                  THAT MATRIX A IS STORED IN A VECTOR FORM.
C           B    - OUTPUT MATRIX (M BY M)  CONTAINING REGRESSION
C                  COEFFICIENTS (B) CORRESPONDING TO THE VALUES OF
C                  INTERCEPTS CONTAINED IN THE OUTPUT MATRIX A.
C           S    - OUTPUT MATRIX (M BY M)  CONTAINING STANDARD ERRORS
C                  OF REGRESSION COEFFICIENTS CORRESPONDING TO THE
C                  COEFFICIENTS CONTAINED IN THE OUTPUT MATRIX B.
C           IER  - 0, NO ERROR.
C                  1, IF NUMBER OF NON-MISSING DATA ELEMENTS FOR J-TH
```

```
C                    VARIABLE IS TWO OR LESS.  IN THIS CASE, STD(J),
C                    SKEW(J), AND CURT(J) ARE SET TO 10**75.  ALL
C                    VALUES OF R, A, B, AND S RELATED TO THIS VARIABLE
C                    ARE ALSO SET TO 10**75.
C                 2, IF VARIANCE OF J-TH VARIABLE IS LESS THAN
C                    10**(-20).  IN THIS CASE, STD(J), SKEW(J), AND
C                    CURT(J) ARE SET TO 10**75.  ALL VALUES OF R, A,
C                    B, AND S RELATED TO THIS VARIABLE ARE ALSO SET TO
C                    10**75.
C
C        REMARKS
C           THIS SUBROUTINE CANNOT DISTINGUISH A BLANK AND A ZERO.
C           THEREFORE, IF A BLANK IS SPECIFIED AS A MISSING DATA CODE IN
C           INPUT CARDS, IT WILL BE TREATED AS 0 (ZERO).
C
C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED
C           NONE
C
C        METHOD
C           LEAST SQUARES REGRESSION LINES AND PRODUCT-MOMENT CORRE-
C           LATION COEFFICIENTS ARE COMPUTED.
C
C     ..................................................................
C
      SUBROUTINE MISR (NO,M,X,CODE,XBAR,STD,SKEW,CURT,R,N,A,B,S,IER)
C
      DIMENSION X(1),CODE(1),XBAR(1),STD(1),SKEW(1),CURT(1),R(1),N(1)
      DIMENSION A(1),B(1),S(1)
C
C        COMPUTE MEANS
C
      IER=0
      L=0
      DO 20 J=1,M
      FN=0.0
      XBAR(J)=0.0
      DO 15 I=1,NO
      L=L+1
      IF(X(L)-CODE(J)) 12, 15, 12
   12 FN=FN+1.0
      XBAR(J)=XBAR(J)+X(L)
   15 CONTINUE
      IF(FN) 16, 16, 17
   16 XBAR(J)=0.0
      GO TO 20
   17 XBAR(J)=XBAR(J)/FN
   20 CONTINUE
C
C        SET-UP WORK AREAS AND TEST WHETHER DATA IS MISSING
C
      L=0
      DO 55 J=1,M
      LJJ=NO*(J-1)
      SKEW(J)=0.0
      CURT(J)=0.0
```

```
      KI=M*(J-1)
      KJ=J-M
      DO 54 I=1,J
      KI=KI+1
      KJ=KJ+M
      SUMX=0.0
      SUMY=0.0
      TI=0.0
      TJ=0.0
      TII=0.0
      TJJ=0.0
      TIJ=0.0
      NIJ=0
      LI=NO*(I-1)
      LJ=LJJ
      L=L+1
      DO 38 K=1,NO
      LI=LI+1
      LJ=LJ+1
      IF(X(LI)-CODE(I)) 30, 38, 30
   30 IF(X(LJ)-CODE(J)) 35, 38, 35
C
C        BOTH DATA ARE PRESENT
C
   35 XX=X(LI)-XBAR(I)
      YY=X(LJ)-XBAR(J)
      TI=TI+XX
      TII=TII+XX**2
      TJ=TJ+YY
      TJJ=TJJ+YY**2
      TIJ=TIJ+XX*YY
      NIJ=NIJ+1
      SUMX=SUMX+X(LI)
      SUMY=SUMY+X(LJ)
      IF(I-J) 38, 37, 37
   37 SKEW(J)=SKEW(J)+YY**3
      CURT(J)=CURT(J)+YY**4
   38 CONTINUE
C
C        COMPUTE SUM OF CROSS-PRODUCTS OF DEVIATIONS
C
      IF(NIJ) 40, 40, 39
   39 FN=NIJ
      R(L)=TIJ-TI*TJ/FN
      N(L)=NIJ
      TII=TII-TI*TI/FN
      TJJ=TJJ-TJ*TJ/FN
C
C        COMPUTE STANDARD DEVIATION, SKEWNESS, AND KURTOSIS
C
   40 IF(I-J) 47, 41, 47
   41 IF(NIJ-2) 42,42,43
   42 IER=1
      R(L)=1.0E75
      A(KI)=1.0E75
```

```
      B(KI)=1.0E75
      S(KI)=1.0E75
      GO TO 45
C
   43 STD(J)=R(L)
      R(L)=1.0
      A(KI)=0.0
      B(KI)=1.0
      S(KI)=0.0
C
      IF(STD(J)-(1.0E-20)) 44,44,46
   44 IER=2
   45 STD(J)=1.0E75
      SKEW(J)=1.0E75
      CURT(J)=1.0E75
      GO TO 55
C
   46 WORK=STD(J)/FN
      SKEW(J)=(SKEW(J)/FN)/(WORK*SQRT(WORK))
      CURT(J)=((CURT(J)/FN)/WORK**2)-3.0
      STD(J)=SQRT(STD(J)/(FN-1.0))
      GO TO 55
C
C        COMPUTE REGRESSION COEFFICIENTS
C
   47 IF(NIJ-2) 48,48,50
   48 IER=1
   49 R(L)=1.0E75
      A(KI)=1.0E75
      B(KI)=1.0E75
      S(KI)=1.0E75
      A(KJ)=1.0E75
      B(KJ)=1.0E75
      S(KJ)=1.0E75
      GO TO 54
C
   50 IF(TII-(1.0E-20)) 52,52,51
   51 IF(TJJ-(1.0E-20)) 52,52,53
   52 IER=2
      GO TO 49
C
   53 SUMX=SUMX/FN
      SUMY=SUMY/FN
      B(KI)=R(L)/TII
      A(KI)=SUMY-B(KI)*SUMX
      B(KJ)=R(L)/TJJ
      A(KJ)=SUMX-B(KJ)*SUMY
C
C        COMPUTE CORRELATION COEFFICIENTS
C
      R(L)=R(L)/(SQRT(TII)*SQRT(TJJ))
C
C        COMPUTE STANDARD ERRORS OF REGRESSION COEFFICIENTS
C
      RR=R(L)**2
```

```
      SUMX=(TJJ-TJJ*RR)/(FN-2)
      S(KI)=SQRT(SUMX/TII)
      SUMY=(TII-TII*RR)/(FN-2)
      S(KJ)=SQRT(SUMY/TJJ)
C
   54 CONTINUE
   55 CONTINUE
C
      RETURN
      END
```

```
C -----------------------------
C     PROGRAM PFOURIER
C -----------------------------
C
C COMPUTE FOURIER COEFFICIENTS FOR
C SAMPLE DATA.
C AND PLOT IN X-Y PLOT FORM
C
 DIMENSION X(301),A(51),B(51)
C OPEN THE FILE
 IF (IOREAD (5,0,0,"THDAT.FIL")) GOTO 999
 IF(IOWRIT(6,2,0,"PFOURIER.FIL")) STOP 8
C READ DATA
 DO 50 I=1,301
 READ (5,100,ENDFILE=60) T,TGT,HR,ERR
100 FORMAT (F5.1,3F8.3)
 X(I)=HR
50 WRITE (1,150)
150 FORMAT (2H *$
60 CONTINUE
 NCOEFF = 50
 CALL    FORIT (X,150,NCOEFF,A,B,IER)
 WRITE (4) "              FOURIER COEFF."
 WRITE (4,180)
180 FORMAT (1H /)
 WRITE (4)"    FREQ      COS      SIN     AMPL    PHASE"
 WRITE (4,180)
 DO 300 J=1,NCOEFF
 AJ=J-1
 FREQ=AJ*(1.0/60.0)
 AMPL=SQRT(A(J)*A(J)+B(J)*B(J))
 PHASE = 57.3*ATAN2(A(J),B(J))
 WRITE (6,190) FREQ, A(J),B(J),AMPL,PHASE
 A(J)=AMPL ; B(J)=FREQ
190 FORMAT (4F8.3,F8.0)
300 CONTINUE
 CALL PLT (B,A,50,0.0,1.0,0.0,10.0)
200 FORMAT (8F8.3)
 IF(IOCLOS(6)) STOP 9
55 IF (IOCLOS(5)) GOTO 999
 STOP
999 WRITE (1) "ERROR"
 END
C
C     ...............................................................
C
C        SUBROUTINE FORIT
C
C        PURPOSE
C           FOURIER ANALYSIS OF A PERIODICALLY TABULATED FUNCTION.
C           COMPUTES THE COEFFICIENTS OF THE DESIRED NUMBER OF TERMS
C           IN THE FOURIER SERIES F(X)=A(0)+SUM(A(K)COS KX+B(K)SIN KX)
C           WHERE K=1,2,...,M TO APPROXIMATE A GIVEN SET OF
C           PERIODICALLY TABULATED VALUES OF A FUNCTION.
C
```

```
C        USAGE
C           CALL FORIT(FNT,N,M,A,B,IER)
C
C        DESCRIPTION OF PARAMETERS
C           FNT-VECTOR OF TABULATED FUNCTION VALUES OF LENGTH 2N+1
C           N  -DEFINES THE INTERVAL SUCH THAT 2N+1 POINTS ARE TAKEN
C               OVER THE INTERVAL (0,2PI). THE SPACING IS THUS 2PI/2N+1
C           M  -MAXIMUM ORDER OF HARMONICS TO BE FITTED
C           A  -RESULTANT VECTOR OF FOURIER COSINE COEFFICIENTS OF
C               LENGTH M+1
C               A SUB 0, A SUB 1,..., A SUB M
C           B  -RESULTANT VECTOR OF FOURIER SINE COEFFICIENTS OF
C               LENGTH M+1
C               B SUB 0, B SUB 1,..., B SUB M
C           IER-RESULTANT ERROR CODE WHERE
C               IER=0  NO ERROR
C               IER=1  N NOT GREATER OR EQUAL TO M
C               IER=2  M LESS THAN 0
C
C        REMARKS
C           M MUST BE GREATER THAN OR EQUAL TO ZERO
C           N MUST BE GREATER THAN OR EQUAL TO M
C           THE FIRST ELEMENT OF VECTOR B IS ZERO IN ALL CASES
C
C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED
C           NONE
C
C        METHOD
C           USES RECURSIVE TECHNIQUE DESCRIBED IN A. RALSTON, H. WILF,
C           'MATHEMATICAL METHODS FOR DIGITAL COMPUTERS', JOHN WILEY
C           AND SONS, NEW YORK, 1960, CHAPTER 24. THE METHOD OF INDEXING
C           THROUGH THE PROCEDURE HAS BEEN MODIFIED TO SIMPLIFY THE
C           COMPUTATION.
C
C     ..................................................................
C
      SUBROUTINE FORIT(FNT,N,M,A,B,IER)
      DIMENSION A(51),B(51),FNT(301)
C
C        CHECK FOR PARAMETER ERRORS
C
      IER=0
   20 IF(M) 30,40,40
   30 IER=2
      RETURN
   40 IF(M-N) 60,60,50
   50 IER=1
      RETURN
C
C        COMPUTE AND PRESET CONSTANTS
C
   60 AN=N
      COEF=2.0/(2.0*AN+1.0)
      CONST=3.141593*COEF
      S1=SIN(CONST)
```

```
      C1=COS(CONST)
      C=1.0
      S=0.0
      J=1
      FNTZ=FNT(1)
   70 U2=0.0
      U1=0.0
      I=2*N+1
C
C        FORM FOURIER COEFFICIENTS RECURSIVELY
C
   75 U0=FNT(I)+2.0*C*U1-U2
      U2=U1
      U1=U0
      I=I-1
      IF(I-1) 80,80,75
   80 A(J)=COEF*(FNTZ+C*U1-U2)
      B(J)=COEF*S*U1
      IF(J-(M+1)) 90,100,100
   90 Q=C1*C-S1*S
      S=C1*S+S1*C
      C=Q
      J=J+1
      GO TO 70
  100 A(1)=A(1)*0.5
      RETURN
      END
```

```
400 CONTINUE
 DO 500 J1=1,51
 J=52-J1
 DO 450 I1=1,50
 I2=52-I1
 IF(IA(I2,J).NE.ISP) GO TO 475
450 CONTINUE
475 IF(MOD((J-1),5).EQ.0) GO TO 490
 WRITE (6,20) (IA(I,J),I=1,I2)
 GO TO 499
490 WRITE (6,30) YLAB(J/5+1),(IA(I,J),I=1,I2)
499 CONTINUE
500 CONTINUE
 WRITE (6,40)
 WRITE (6,50) (XLAB(I),I=1,11,2)
 RETURN
 END
```

```
C
C     ...............................................................
C
C        SUBROUTINE MULTR
C
C        PURPOSE
C           PERFORM A MULTIPLE LINEAR REGRESSION ANALYSIS FOR A
C           DEPENDENT VARIABLE AND A SET OF INDEPENDENT VARIABLES.  THIS
C           SUBROUTINE IS NORMALLY USED IN THE PERFORMANCE OF MULTIPLE
C           AND POLYNOMIAL REGRESSION ANALYSES.
C
C        USAGE
C           CALL MULTR (N,K,XBAR,STD,D,RX,RY,ISAVE,B,SB,T,ANS)
C
C        DESCRIPTION OF PARAMETERS
C           N     - NUMBER OF OBSERVATIONS.
C           K     - NUMBER OF INDEPENDENT VARIABLES IN THIS REGRESSION.
C           XBAR  - INPUT VECTOR OF LENGTH M CONTAINING MEANS OF ALL
C                   VARIABLES. M IS NUMBER OF VARIABLES IN OBSERVATIONS.
C           STD   - INPUT VECTOR OF LENGTH M CONTAINING STANDARD DEVI-
C                   ATIONS OF ALL VARIABLES.
C           D     - INPUT VECTOR OF LENGTH M CONTAINING THE DIAGONAL OF
C                   THE MATRIX OF SUMS OF CROSS-PRODUCTS OF DEVIATIONS
C                   FROM MEANS FOR ALL VARIABLES.
C           RX    - INPUT MATRIX (K X K) CONTAINING THE INVERSE OF
C                   INTERCORRELATIONS AMONG INDEPENDENT VARIABLES.
C           RY    - INPUT VECTOR OF LENGTH K CONTAINING INTERCORRELA-
C                   TIONS OF INDEPENDENT VARIABLES WITH DEPENDENT
C                   VARIABLE.
C           ISAVE - INPUT VECTOR OF LENGTH K+1 CONTAINING SUBSCRIPTS OF
C                   INDEPENDENT VARIABLES IN ASCENDING ORDER.  THE
C                   SUBSCRIPT OF THE DEPENDENT VARIABLE IS STORED IN
C                   THE LAST, K+1, POSITION.
C           B     - OUTPUT VECTOR OF LENGTH K CONTAINING REGRESSION
C                   COEFFICIENTS.
C           SB    - OUTPUT VECTOR OF LENGTH K CONTAINING STANDARD
C                   DEVIATIONS OF REGRESSION COEFFICIENTS.
C           T     - OUTPUT VECTOR OF LENGTH K CONTAINING T-VALUES.
C           ANS   - OUTPUT VECTOR OF LENGTH 10 CONTAINING THE FOLLOWING
C                   INFORMATION..
C                   ANS(1)  INTERCEPT
C                   ANS(2)  MULTIPLE CORRELATION COEFFICIENT
C                   ANS(3)  STANDARD ERROR OF ESTIMATE
C                   ANS(4)  SUM OF SQUARES ATTRIBUTABLE TO REGRES-
C                           SION (SSAR)
C                   ANS(5)  DEGREES OF FREEDOM ASSOCIATED WITH SSAR
C                   ANS(6)  MEAN SQUARE OF SSAR
C                   ANS(7)  SUM OF SQUARES OF DEVIATIONS FROM REGRES-
C                           SION (SSDR)
C                   ANS(8)  DEGREES OF FREEDOM ASSOCIATED WITH SSDR
C                   ANS(9)  MEAN SQUARE OF SSDR
C                   ANS(10) F-VALUE
C
C        REMARKS
C           N MUST BE GREATER THAN K+1.
```

```
C
C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED
C           NONE
C
C        METHOD
C           THE GAUSS-JORDAN METHOD IS USED IN THE SOLUTION OF THE
C           NORMAL EQUATIONS.  REFER TO W. W. COOLEY AND P. R. LOHNES,
C           'MULTIVARIATE PROCEDURES FOR THE BEHAVIORAL SCIENCES',
C           JOHN WILEY AND SONS, 1962, CHAPTER 3, AND B. OSTLE,
C           'STATISTICS IN RESEARCH', THE IOWA STATE COLLEGE PRESS,
C           1954, CHAPTER 8.
C
C     ..................................................................
C
      SUBROUTINE MULTR (N,K,XBAR,STD,D,RX,RY,ISAVE,B,SB,T,ANS)
      DIMENSION XBAR(1),STD(1),D(1),RX(1),RY(1),ISAVE(1),B(1),SB(1),
     1          T(1),ANS(1)
C
C        ...............................................................
C
C        IF A DOUBLE PRECISION VERSION OF THIS ROUTINE IS DESIRED, THE
C        C IN COLUMN 1 SHOULD BE REMOVED FROM THE DOUBLE PRECISION
C        STATEMENT WHICH FOLLOWS.
C
C     DOUBLE PRECISION XBAR,STD,D,RX,RY,B,SB,T,ANS,RM,BO,SSAR,SSDR,SY,
C    1                 FN,FK,SSARM,SSDRM,F
C
C        THE C MUST ALSO BE REMOVED FROM DOUBLE PRECISION STATEMENTS
C        APPEARING IN OTHER ROUTINES USED IN CONJUNCTION WITH THIS
C        ROUTINE.
C
C        THE DOUBLE PRECISION VERSION OF THIS SUBROUTINE MUST ALSO
C        CONTAIN DOUBLE PRECISION FORTRAN FUNCTIONS.  SQRT AND ABS IN
C        STATEMENTS 122, 125, AND 135 MUST BE CHANGED TO DSQRT AND DABS.
C
C        ...............................................................
C
      MM=K+1
C
C        BETA WEIGHTS
C
      DO 100 J=1,K
  100 B(J)=0.0
      DO 110 J=1,K
      L1=K*(J-1)
      DO 110 I=1,K
      L=L1+I
  110 B(J)=B(J)+RY(I)*RX(L)
      RM=0.0
      BO=0.0
      L1=ISAVE(MM)
C
C        COEFFICIENT OF DETERMINATION
C
      DO 120 I=1,K
```

```
      RM=RM+B(I)*RY(I)
C
C        REGRESSION COEFFICIENTS
C
      L=ISAVE(I)
      B(I)=B(I)*(STD(L1)/STD(L))
C
C        INTERCEPT
C
  120 BO=BO+B(I)*XBAR(L)
      BO=XBAR(L1)-BO
C
C        SUM OF SQUARES ATTRIBUTABLE TO REGRESSION
C
      SSAR=RM*D(L1)
C
C        MULTIPLE CORRELATION COEFFICIENT
C
  122 RM= SQRT( ABS(RM))
C
C        SUM OF SQUARES OF DEVIATIONS FROM REGRESSION
C
      SSDR=D(L1)-SSAR
C
C        VARIANCE OF ESTIMATE
C
      FN=N-K-1
      SY=SSDR/FN
C
C        STANDARD DEVIATIONS OF REGRESSION COEFFICIENTS
C
      DO 130 J=1,K
      L1=K*(J-1)+J
      L=ISAVE(J)
  125 SB(J)= SQRT( ABS((RX(L1)/D(L))*SY))
C
C        COMPUTED T-VALUES
C
  130 T(J)=B(J)/SB(J)
C
C        STANDARD ERROR OF ESTIMATE
C
  135 SY= SQRT( ABS(SY))
C
C        F VALUE
C
      FK=K
      SSARM=SSAR/FK
      SSDRM=SSDR/FN
      F=SSARM/SSDRM
C
      ANS(1)=BO
      ANS(2)=RM
      ANS(3)=SY
      ANS(4)=SSAR
```

```
ANS(5)=FK
ANS(6)=SSARM
ANS(7)=SSDR
ANS(8)=FN
ANS(9)=SSDRM
ANS(10)=F
RETURN
END
```

```
C     ..................................................................
C
C        SUBROUTINE CROSS
C
C        PURPOSE
C           TO FIND THE CROSSCOVARIANCES OF SERIES A WITH SERIES B
C           (WHICH LEADS AND LAGS A).
C
C        USAGE
C           CALL CROSS (A,B,N,L,R,S)
C
C        DESCRIPTION OF PARAMETERS
C           A     - INPUT VECTOR OF LENGTH N CONTAINING FIRST TIME
C                   SERIES.
C           B     - INPUT VECTOR OF LENGTH N CONTAINING SECOND TIME
C                   SERIES.
C           N     - LENGTH OF SERIES A AND B.
C           L     - CROSSCOVARIANCE IS CALCULATED FOR LAGS AND LEADS OF
C                   0, 1, 2,..., L-1.
C           R     - OUTPUT VECTOR OF LENGTH L CONTAINING CROSSCOVARI-
C                   ANCES OF A WITH B, WHERE B LAGS A.
C           S     - OUTPUT VECTOR OF LENGTH L CONTAINING CROSSCOVARI-
C                   ANCES OF A WITH B, WHERE B LEADS A.
C
C        REMARKS
C           N MUST BE GREATER THAN L.  IF NOT, R(1) AND S(1) ARE SET TO
C           ZERO AND RETURN IS MADE TO THE CALLING PROGRAM.
C
C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED
C           NONE
C
C        METHOD
C           DESCRIBED IN R.B. BLACKMAN AND J.W. TUKEY, 'THE MEASURMENT
```

```
C        OF POWER SPECTRA', DOVER PUBLICATIONS INC., NEW YORK, 1959.
C
C     ...............................................................
C
      SUBROUTINE CROSS (A,B,N,L,R,S)
      DIMENSION A(1),B(1),R(1),S(1)
C
C     CALCULATE AVERAGES OF SERIES A AND B
C
      FN=N
      AVERA=0.0
      AVERB=0.0
      IF(N-L)50,50,100
   50 R(1)=0.0
      S(1)=0.0
      RETURN
  100 DO 110 I=1,N
      AVERA=AVERA+A(I)
  110 AVERB=AVERB+B(I)
      AVERA=AVERA/FN
      AVERB=AVERB/FN
C
C     CALCULATE CROSSCOVARIANCES OF SERIES A AND B
C
      DO 130 J=1,L
      NJ=N-J+1
      SUMR=0.0
      SUMS=0.0
      DO 120 I=1,NJ
      IJ=I+J-1
      SUMR=SUMR+(A(I)-AVERA)*(B(IJ)-AVERB)
  120 SUMS=SUMS+(A(IJ)-AVERA)*(B(I)-AVERB)
      FNJ=NJ
      R(J)=SUMR/FNJ
  130 S(J)=SUMS/FNJ
      RETURN
      END
```

```
100 FORMAT (F5.1,3F8.3)
 END
C
C
C
 SUBROUTINE PLT (X,Y,N,XMIN,XMAX,YMIN,YMAX)
 DIMENSION X(201),Y(201),IA(51,51),XLAB(11),YLAB(11)
 DATA ISP,IPR,IS/1H ,1H* ,1HX/
20 FORMAT (10X,1H+,80A1)
30 FORMAT (1X,E9.2,1H-,80A1)
40 FORMAT (10X,2H+I,5(10H+++++++++I))
50 FORMAT (6X,6E10.2)
 WRITE (1) "START PLOT"
 XSP =50.
 YSP=XSP
 DX=XSP/(XMAX-XMIN)
 DY=YSP/(YMAX-YMIN)
 DO 200 I=1,51
 DO 200 J=1,51
200 IA(I,J)=ISP
 DO 300 K=1,N
 I=DX*(X(K)-XMIN)+1.5
 J=DY*(Y(K)-YMIN)+1.5
 IA(I,J)=IPR
300 CONTINUE
 DO 400 I=1,51,5
 J=I/5+1
 XLAB(J)=(I-1)/DX+XMIN
 XLAB(1)=XLAB(1)
 YLAB(J)=(I-1)/DY+YMIN
400 CONTINUE
 DO 500 J1=1,51
 J=52-J1
 DO 450 I1=1,50
 I2=52-I1
 IF(IA(I2,J).NE.ISP) GO TO 475
450 CONTINUE
475 IF(MOD((J-1),5).EQ.0) GO TO 490
 WRITE (4,20) (IA(I,J),I=1,I2)
 GO TO 499
490 WRITE (4,30) YLAB(J/5+1),(IA(I,J),I=1,I2)
499 CONTINUE
500 CONTINUE
 WRITE (4,40)
 WRITE (4,50) (XLAB(I),I=1,11,2)
 RETURN
 END
 U0 = .2611597647
 AK=37.0
 XBAR=0.0
 XSQU=0.0
 COUNT = 10000.0
 C1=0.0
 C9=0.0
 C4=0.0
```

APPENDIX C

COMMENT ON STATISTICAL MODELS OF PERFORMANCE

The goal of any performance measurement system is to capture the complexity of the real world while providing concise metrics. When there exists a lawful relationship between pilot skill level and the performance variables associated with the pilot-aircraft system, response surface modeling is a possible alternative. If the measurement variables can be designed so as to represent a quantifiable continuum on the skill axis, multivariate statistical techniques become a powerful tool for the development of response surface models.

Performance models are functions composed of several variables or measures useful for higher-level decision making or automated program control. Simply put, these multivariate models could simultaneously combine information more effectively than humans could. The simplest models or functions take the linear form:

$$Y = b_1x_1 + b_2x_2 + b_3x_3 + \ldots + b_mx_m + E$$

The Y term is usually the effect that the equation is trying to enumerate and the b's are weighting coefficients derived by some multivariate analyses using several samples of m number of variables or x's. E represents the error term of variance not accounted for by the model.

Assuming existence of a lawful relationship between m variables, there is . m-dimensional surface which describes this relationship, to any error level, with a polynomial of high enough order. Of course, very high order polynominals are more difficult to interpret and implement.

A complex version of this method would be to develop and explore a surface model of <u>all</u> the measures taken from the system, before having removed some of them from the model. In the process of doing this a greater insight into the inter-relationship of all measures may be gained, (thus reducing the error of throwing out information unnecessarily) but it could be an extensive task only recommended if the simpler models are not satisfactory.

If the relationships between all the measures can be reasonable determined in this manner, then sample size becomes of lesser importance and the extra time expended in the surface modeling phase is repaid several times over by the abbreviated data gathering segment of the study. In otherwords, time is probably better spent on adequately describing the information contained in the data base, rather than accumulating subjects while attempting to render significance to some small effects in the data.

The performance measurement modeling approach becomes a simple concept with great user validity. Multivariate statistical techniques are used to characterize the change in system performance variables as they change with the experience and training of pilots. Since the performance measures can be taken at various well defined stages in training while performing the

actual task of interest, the performance measurement model is operationally defined in the field setting. Any error term associated with inter-pilot differences can be further reduced with the addition of variables to account for pilot characteristics, for example: age, sex or educational experience.

This appendix will discuss the theoretical and practical aspects of statistical performance measurement modeling by way of the following steps.

1. Reduction of Dimensionality.
2. Development and Study of Multivariate Relationships.
3. Statistical Issues in the Assessment of Multivariate Models.

Several statistical procedures will be described and their respective FORTRAN IV programs will be included in this Appendix. Some algorithms described in the text will not be included in the appendix as they are commonly found in most statistical libraries or only simple modifications to existing analysis routines.

The discussion in this section relates entirely to multivariate statistical techniques which have been developed over the years and have been used successfully on more than one occasion. This is not meant to be a tutorial on data analytic methods in the broad sense. There will be no attention given to analysis of time series, and the performance measurement development procedures used preclude the discussion of any pattern recognition or multidimensional classification techniques since group membership is assigned before data collection. The following discussion presumes, in fact, that pilots can be divided into groups representing various skill levels and a set of summary performance variables have been collected on each while performing precisely the same task. The task should be a segment of a maneuver or procedure under quantifiable conditions (preferably the same) with well defined initial and terminal conditions.

The following methods seek to develop multivariate performance measurement models that are maximally sensitive to the differences, or discriminate between skill groups.

REDUCTION OF DIMENSIONALITY

Ideally, the discriminant system should be composed of a small set of entirely orthogonal or independent variables which, when combined into a single function, will describe 100 percent of the variance contained in the data.

Unfortunately, few situations are encountered which provide variables so well behaved. Even with the best of discriminant functions, there usually remains a small overlap or region of indecision between classifications. It becomes possible only to reduce this unexplained variance with the best selection of variables.

In addition, rarely are variables totally orthogonal to each other; they will be correlated to some degree. This means that the relative

contribution of each variable cannot be directly assessed, but can be inferred by careful examination of both univariate and multivariate statistical procedures.

Aircrew data will contain these problems. Many aircraft control variables are highly correlated, and it is expected that some overlap will occur in the control activity that they describe. With these difficulties in mind, this section will discuss several procedures available to select variables for the discriminant system.

Data Editing. So far, we have only considered the elimination of variables from the discriminant system, but the number of samples analyzed represents another dimension that may suffer reductions prior to computation the scores. The controlled reduction of variables in the discriminant model has some beneficial effects, however, the exclusion of samples reduces the information from which the system will be developed. It is highly possible that outliers (a few unusual performances) may adversely bias the entire discriminant function and an improvement in fit may result only from their removal before analysis.

Rules have been proposed for automatically rejecting outliers. Our experience has shown a significant improvement in discrimination sometimes results from the removal of a very small percentage of data. The simplest of these rules might be to avoid any sample which contains a variable with a value more than 3.5 standard deviations away from the mean of all samples. However, this value can be changed, based on examination of the data.

This is not as risky a procedure as some would suggest, because outliers will not be ignored in the final implementation of the scoring system. To avoid applying the discriminant function to performances outside of the range of data used to create it (a common error of many projections), outliers can be detected by the same criteria, and nominated for examination rather than being subjected to the discriminant scoring. Thus, outliers will always receive special attention while the remaining performances can be scored with a more optimal discriminator.

Performance measures can be automatically edited, and the results tabulated for further application, prior to any statistical processing. This procedure will result in several other fringe benefits, as discussed in the following sections.

Univariate Selection. In some cases variables can be selected or rejected on the basis of their individual ability to classify performance. Immediately after data editing, it is a time conserving step to remove those variables which have a trivial relationship to skill change or group membership. This eliminates wasteful over processing of easily discernable, useless information. Since the final set of variables to be used represents a small fraction of the initial set, it is a safe strategy to select a reasonably large starting set without regard to the quality of individual distributions or other sophisticated multivariate considerations.

The rationale behind this strategy is straight forward. Under the best of conditions, given the variables were to be orthogonal, the portion of variance each accounted for in the final multivariate model would be equal to the respective portion they accounted for individually in a univariate model. As their inter-correlations increase, the actual portion of the multivariate variance they account for decreases from what the univariate models would suggest.

Further, it is suggested to perform the <u>initial</u> screening of variables before any data are transformed for the following reasons: the criterion significance level so low that any small difference between groups will pass the test. The distributions of the variates within each group may be an issue for this part of the procedure. Should the sample sizes be large ($N > 20$) a few factors will work together to improve the robustness of most univariate statistical procedures for these purposes. A large number of samples provides a stable estimate of the population. The extreme tails of any of the distributions have been clipped by the data editing procedure to improve their normality. Essentially it is at the tails or higher percentiles that most of the tests become sensitive to fluctuations in the distributions. We have effectively avoided this problem by only requiring a minimally significant effect.

Should the sample sizes be small, the univariate distributions are likely to be ill-conditioned and normal parametric statistics are likely to provide less reliable information. In this case, some distribution free or nonparametric analyses may be more reliable. Non-parametric techniques require little or no degrees of freedom which makes them highly suited for situations with very few observations.

The "Tukey Quick Test of Location" (Bruning and Kintz, 1977; Hays, 1973; Neave, 1979) can be used to assess the significance of any difference embodied in a measure across skill groups. This test simply requires the grand minimum and grand maximum for the measure in question to <u>not</u> be contained in the same skill group. In other words, one group must contain the minimum for that measure for all observations and the other group must contain the maximum. The Tukey Test therefore, may have critical limitations when applied to certain classes of variables; specifically those that could assume a zero value in more than one skill group. A measure must be subjected to another test if it fails this test. All the confirmatory occurrences are then counted and "Tukey Counts" are assigned significance levels similar to t-tests. Experimenter judgment is used to assign an acceptance or screening level for selecting the most significant measures.

Our experience has shown that when univariate tests for all the variables are examined many variables are practically invariant in relation to the dependent variable. Removal of these trivial variables, usually leaves a sufficient number of 'active' variables which relate to the dependent variable in significant ways.

For example, if a t-test with a t equal to 0.70 is applied to each of the approximately 300 performance measures, it may leave 60 significant

variables. Considering the final set will contain some 15 variables, 60 is hardly a skimpy starting point for more careful examination.

Multivariate Selection Schemes. Once outliers and obviously trivial measures have been removed from the data base, the remaining performance measures will be examined with a much finer multivariate screening procedure. First, in the same vein as removing trivial variables, it is important also to remove very highly correlated dependent variables. Secondly, the final set of variables in the discriminant function will be determined by step down discriminant procedure which removes those variables which account for the least variance in the discrimination. These two procedures are described in the paragraphs that follow.

Elimination from the analysis is suggested for one measure from a near-perfectly correlated pair, since they practically describe the same aspect of the function. By near-perfectly correlated, we mean pairs of variables with correlations greater than .98. This criterion can be adjusted for pragmatic consideration related specifically to prior knowledge about those specific return characteristics or the effect that the variables removal may have on the power of the discriminant function. To insure the least impact on the function, only the variable in the pair which correlates most with every other variable in the system will be removed. The least information will be lost as a result of this strategy, since only the variables that have the most discriminant information in common with all the other variables shall be removed. Another benefit of this procedure is to improve the condition of the correlation matrix for the discriminant analysis; the importance of this is described in the subsection called 'discriminant analysis'.

With all of the data preconditioning complete, the step-down discriminant analysis is performed. Discriminant analysis is iteratively performed, removing measures which contribute least to the model. Communality (Cooley and Lohnes, 1971) is therefore the criterion for retention in the model. Control of this process is done by varying the minimum communality allowed, and by initial inclusion of measures at the start of the process. The software also is designed to force the inclusion of variables into the function at the discretion of the analyst.

DEVELOPMENT AND STUDY OF MULTIVARIATE RELATIONSHIPS

Discriminant Analysis. Discriminant analysis is useful for reducing a multivariate problem to a linear function of the variates which maximizes the difference between populations:

$$Y=b_1x_1+b_2x_2+b_3x_3+\ldots+b_nx_n.$$

In review, multiple discriminant analysis projects data points from their initial measurement space into a suitable subspace. This subspace is univariate and usually referred to as discriminant space. The discriminant model determines those components which best separate the groups in measurement space, and weights them to maximize this difference in discriminant space. The geometric interpretation of discriminant analysis can be seen, for the case of two groups and two variates, in Figure 1.

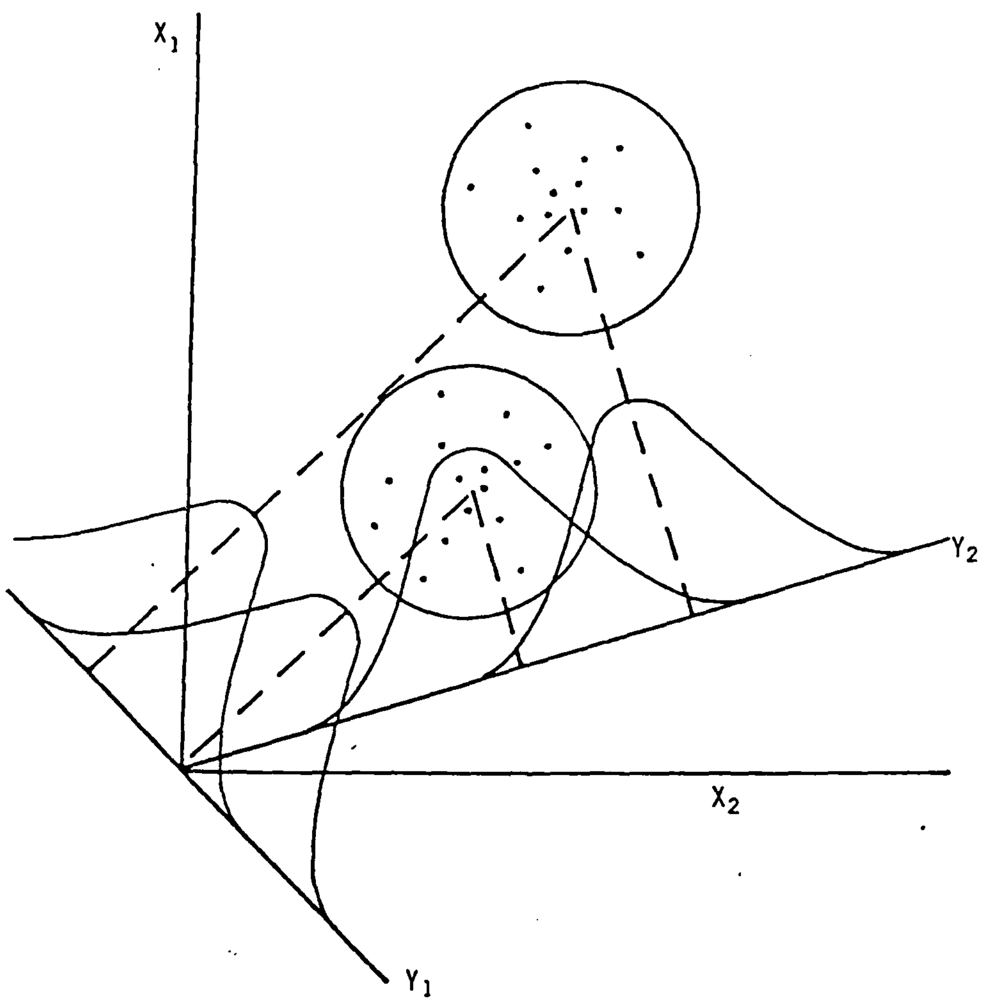


Figure 1. Different Linear Combinations for Two Variables.

In Figure 1 can be seen two partially overlapping bivariate-normal scatter diagrams projected onto a new axis ¶. The two new overlapping distributions represent the two groups projected onto an arbitrary discriminant axis. The degree of overlap can be manipulated by varying the discriminant coefficients used to transform the multivariate points onto the resulting discriminant axis. The objective of discriminant analysis is to find a set of coefficients which minimize this projected overlap for two or more groups consisting of many normally distributed variables.

It makes no difference to the formal logic of the discriminant model whether the variates in measurement space are the dependent variables and the discriminant function is the independent predictor vector, or the groups consist of independent treatment variables and group membership is the dependent variable vectors. For example, in the case of performance measurement development (Vreuls, D., Wooldridge, L., et al., 1977) the groups represented relative skill levels of pilots composed of dependent and uncontrolled performance measures and the group membership was the only controlled independent variable. In this case the return characteristics will be the dependent or uncontrolled variables and the group membership will be an 'a posteriori' controlled independent variable.

<u>Violation of Statistical Assumptions</u>. Often, pilot performance measures are not collected with carefully designed experimental conditions, and the measures are not specifically intended to be independent variables, nor are they expected to have normal (Gaussian) distributions.

Non-normal distributions can be detected and rectified by using the appropriate standard, normalizing transformations. In particular, error measures usually have Poisson or Weibull Distributions (Gibra, 1973). Unequal dispersions are quite common and can also be favorably corrected with square-root or logarithmic transformations. Care will be taken so as not to unnecessarily degrade any discrimination by using a transformation on an obviously non-normal distribution. It is our goal to provide the most stable and optimal discrimination rather than to satisfy rigorous statistical principals. In some cases, two groups with very different centroids and narrow dispersions discriminate more effectively than if there dispersions were flattened to satisfy arbitrary requirements of normalcy. There is a tradeoff to be made between functional utility and computational complexity and rigorous stastistical theory. It also may be possible to suggest two variable transmutations which would show important interactions between performance measures.

Correlations between the dependent variables are a more stubborn and subtle problem which can dramatically effect the results of a multivariate analysis. The correlations usually are not a result of any causal relationship between the predictor variables.

When the predictor variables are correlated with each other, the intercorrelation matrix will have non-zero correlations in the off-diagonal positions. A hypothetical intercorrelation matrix for the discriminant situation appears in Table 1. This table can be broken into three parts: (1) the predictor matrix of correlations among each predictor variable and every other predictor variable; (2) the diagonal of the matrix which is the correlation of each predictor variable to itself; and, (3) the group membership row vector of correlations between each predictor variable and group membership.

Note that since each predictor variable correlates perfectly and positively with itself, the diagonal values are all one. Since the off-diagonal elements are not zero, the matrix is said to be ill-conditioned and the original experimental design is classified as non-orthogonal.

A multivariate least-squares regression, as well as a multivariate discriminant analysis, will both suffer from similar failings when applied to data characterized by an ill-conditioned intercorrelations matrix. Indeed, there is a popular contention that the regression and discriminant models should always find the same solution for any given data. It has been suggested that any notable differences are due to computational problems. The author feels that the regression criterion behave somewhat

$$(Y-Xb)'(Y-Xb)\,\Big|_{\text{Min}}$$

and the discriminant criterion

$$\frac{b'Ab}{b'Wb}\,\Big|_{\text{Max}}$$

differently under adverse conditions. Should the groups have significantly different shapes, be non-normally distributed or be non-orthogonal, the two criteria may result in quite different discriminant models. Until this relationship is mathematically or practically proven, we reserve the right to discuss the two methods as different, yet analogous, procedures. For the time being, it is this analogy that is of critical importance.

Users of either analysis technique, McDonald and Schwing (1973) for instance, have noted certain instabilities in the resulting ordinary coefficients when analyzing non-orthogonal systems. Some coefficients have extreme magnitudes or incorrect signs resulting in linear functions that respond unsatisfactorily when supplied with new data. This erratic behavior was also noted in the Vreuls and Wooldridge (1976) performance measurement study when new pilots were obviously misclassified by the normal discriminant functions.

TABLE 1

EXAMPLE INTERCORRELATION MATRIX FROM CHARLES SIMON'S (1975) REPORT

		Predictor Variables			Performance
		X1	X2	X3	(Y)
Predictor Variables	X1	1.00	0.145	0.352	0.674
	X2	0.145	1.00	0.022	0.532
	X3	0.352	0.022	1.00	0.348

Hoerl and Kennard (1970) suggested adding a small positive quantity, k, to the unit diagonal of the intercorrelation matrix, X'X, of the predictor variables in regression analysis. The conventional least-squares fit is done using this new matrix to produce what are called 'ridge' coefficients. The standard regression

$$b = (X'X)^{-1}X'Y$$

becomes

$$b = (X'X + kI)^{-1}X'Y$$

where I is the unit diagonal matrix.

The 'ridge' comes from the fact, that as k increases, the variance error decreases more rapidly than the bias error increases. As can be seen in Figure 2, for some value of k, the sum of the bias error and variance error (the mean-square error) is minimized and smaller than it would be for the conventional coefficients. Although the 'ridge' can be mathematically demonstrated to exist, little success has been made in calculating a specific value of k that minimizes the mean-square error. Lindley and Smith (1972); Mallows (1973); and Farebrother (1975), to name a few, all suggest various mathematical criteria for selecting a value of k which would improve the set of coefficients without unduly biasing the estimate.

Hoerl and Kennard did not feel that a mathematical solution for selecting the best k was justified. They proposed visual inspection of the 'ridge' trace. Figure 3 is an example 'ridge' trace. This plots the change in regression coefficients various values of k between 0 and 1.

The following conditions should be looked for when selecting the value of k (in lieu of a mathematical formula):

1. The beta values and particularly their orders of magnitudes have begun to stabilize.

2. The coefficients no longer have unrealistically large absolute values.

3. The coefficients with logically incorrect signs are approaching or have reached the proper sign.

4. The residual sum-of-squares is not unreasonably inflated.

5. The ridge trace (representing the mean-square error) is smaller than the unbiased least-squares variance.

Mathematical Rationale for 'Ridge' Regression. Hoerl and Kennard (1970a) were able to mathematically demonstrate the existence of the ridge estimator for regression by calculating the expected value of the squared distance between $\hat{\beta}^*$ and β. β is the vector of the true regression coefficients and $\hat{\beta}^*$ is the 'ridge' estimates of β. The reader is directed to the popular Hoerl and Kennard paper for the derivation of expected value function $E[L^2(k)]$. Let us only say that an existence theorem can demostrate that,

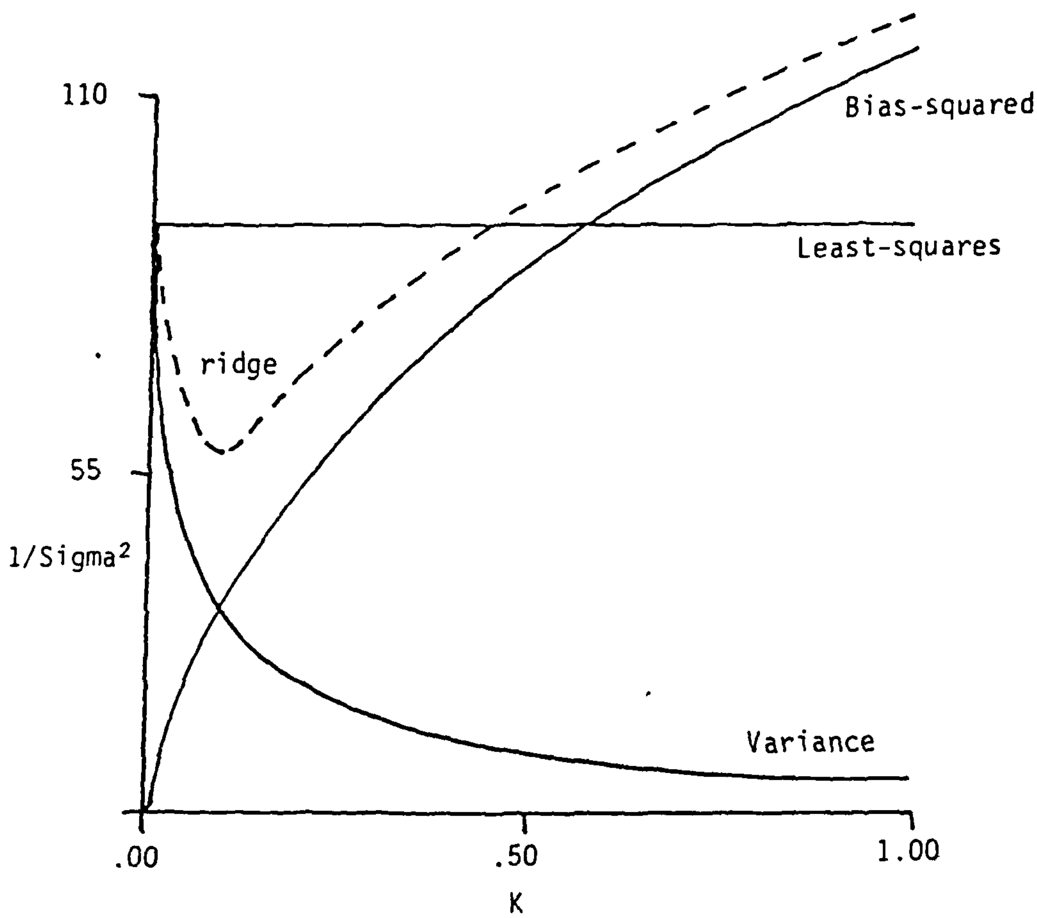


Figure 2. Example Ridge Regression Mean-Square-Error Functions.

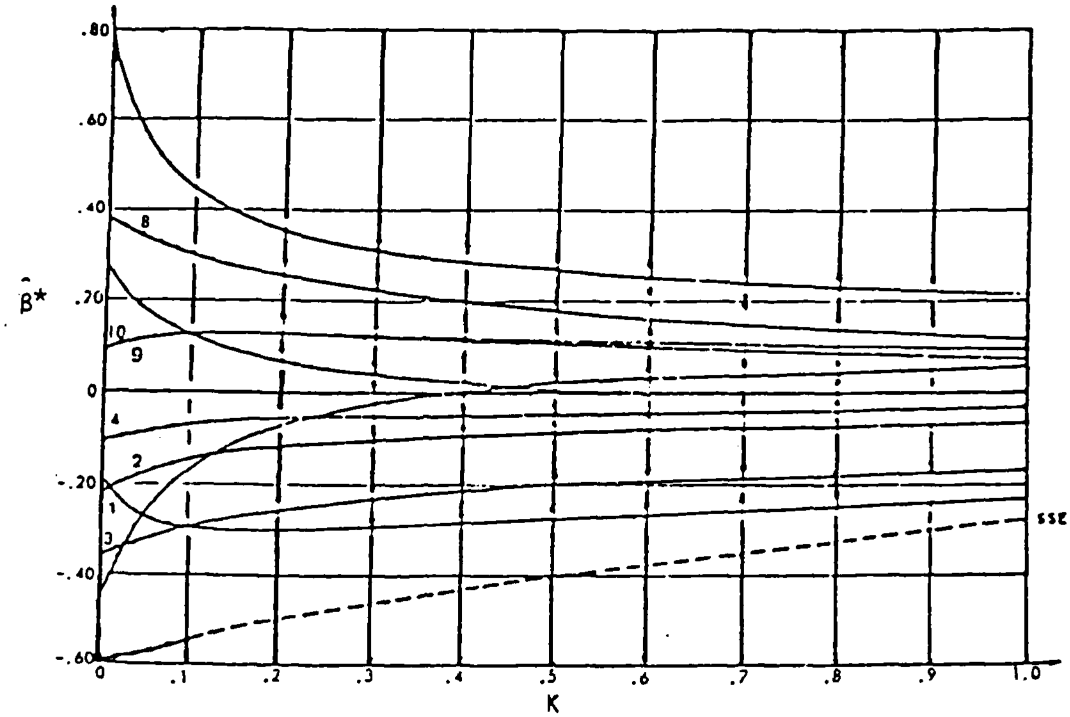


Figure 3. Example Coefficient Trace (Hoerl and Kennard).

although the derivative of E $[L^2(K)]$ is positive as K approaches , there always exists a K>0 such that dE $[L^2(K)]$ / dK<0. This describes the property of E $[L^2(K)]$ of always going through a minimum as K goes from 0 to . Appropriate values of K have been looked for by various solutions to the first derivative of E $[L^2(K)]$.

<u>'Ridge' Discriminant Analysis.</u> Development of a similar expected value function for a 'ridge' discriminant analog has never been done to our knowledge. The best we can offer is an intuitive discussion of the existence of the 'ridge' discriminant adjustment:

In review, discriminant analysis attempts to maximize the criterion,

$$\lambda = \frac{b'Ab}{b'Wb}.$$

Finding the smallest eigen value of W to maximize λ is reasonably comparable to finding the largest eigen value of $W^{-1}A$ used to solve $(W^{-1}A- I)b=0$. The smallest eigen value of W, λ_1, would be very 'small' if two or more experimental variables were correlated. Then if b were chosen to correspond to λ_1 of W then,

$$b'Wb=b'(Wb)=b'(\lambda_1 b)=\lambda_1 b'b=\lambda_1,$$

since $b'b\underline{\Delta}1$. Since λ_1 is 'small', $\frac{b'Ab}{\lambda_1}$ is very 'large'. Undesirably, the solution would disregard or be insensitive to the values found in A. Somewhat larger eigen values of W would be more desirable. One way to force the eigen values to be larger is to replace W with W+IK, where K is a small scalar. Then,

$$b'(R+IK)b=\lambda_1+K.$$

Now, regardless of how small the smallest eigenvector of W is, the size of $_1$+K can be no smaller than K. K being of reasonable size, $\frac{b'Ab}{\lambda_1+K}$ would not be as large as $\frac{b'Ab}{\lambda_1}$. In maximizing $\frac{b'Ab}{b'(W+IK)b}$, A would determine more of b's direction than before. Thus, a very small bias, K, would have the beneficial result of improving the sensitivity of to values found in A. Again, as in 'ridge' regression, there is a trade-off between bias error and variance. It has been demonstrated that the adjusted discriminant function has similar minimizable properties as in the case of least-squares regression. A simulation would also be expected to demonstrate these properties.

Predictably, one characteristic of the 'ridge' discriminant analysis will be quite different from that of the regression version. Usually, the regression case requires only very small values of K, much smaller than 1, usually less than 0.1, to minimize E $[L^2(K)]$. The mathematics of the situation suggests that the discriminant analysis may require adjustments much larger than those typical for regression.

As was previously discussed, several researchers, when using 'ridge' regression, have chosen values for k based on a wide variety of mathematical or pragmatic criteria. The lack of a rigorous mathematical tool for choosing the value of k for discriminant analysis required the development of a Monte Carlo simulation to demonstrate the effectiveness of this new 'ridge' procedure (Bittner 1974).

Using simulated data, the percentage of classification error could be calculated for various values of k. Thus providing an immediate indication of the relative improvement in the discriminant function associated with each value of k. This simulation technique, on the other hand, can provide pseudo-empirical evidence for the selection of a near-optimal discriminant function for any sample simply by finding the minimum classification error as the value of k increases from zero. Rather than base a decision on a highly contestable mathematical assumption, a data simulation should be incorporated as the criteria whenever using a 'ridge' analysis. The specifics of the Monte Carlo simulation will be more fully described in the subsection entitled·'Reliability Test'. Figure 4 shows the simulated misclassifications for a discriminant system developed for the air force (Wooldridge, 1980) over several values of k, note the similarity to the hypothesized ridge in Figure 2. The appropriate value of k is very apparent. Figure shows the functional flow of a suggested 'ridge' discriminant analysis.

Distribution Free Discriminant Analysis. Nonparametric techniques overcome some of problems associated with ill-conditioned distributions or a large number of measures to sample ratio. The resulting discriminant systems would not require the additional variable transformation step associated with the parametric models. The following paragraphs discuss three possible techniques.

The term unit-scaling is used to describe a model using the unit weighting technique with full excursion scaling as opposed to unit normal scaling. The predictor variables are scaled by their sample range after they have been translated by the subtraction of the sample minimum. Thence, the predictor variables are relocated as well as scaled. Although not mathematically required, the translation is performed for convenience of interpretation. The resulting model values for n variances will lie between zero and n. For a set of n raw measures the unit scaling equation would take the form:

$$Y = \frac{a_1x_1 - Min_1}{range_1} + \frac{a_2x_2 - Min_2}{range_2} + \cdots + \frac{a_nx_n - Min_n}{range_n}$$

where a is the unit direction variable, Min is the sample minimum and range is the sample maximum minus the minimum.

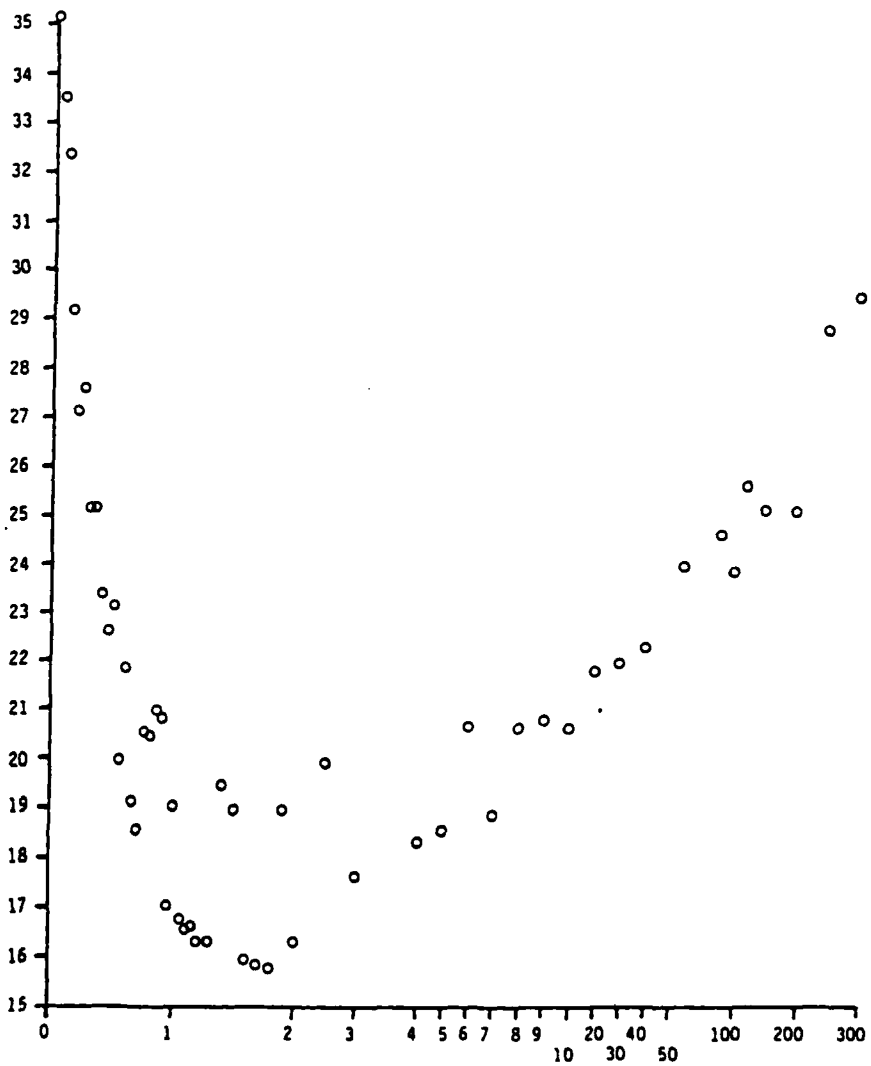


Figure 4. Total Percent Misclassifications for Increasing Values of K for the Instrument Flight Maneuvers Example.

The sample location provides no useful information, so that for either unit weighting or unit scaling, the predictor variable could be relocated in the same manner. Disregarding the sign variable and the relocation, the comparable coefficients become "+" respectively. The primary difference

$$\frac{1}{sigma_m} \text{ and } \frac{1}{range_m}$$

between the standard deviation and the range is that the latter is totally unaffected by the shape of the sample distribution. This feature provides more latitude in the application of the equal weighting scheme.

Given a set of scores for each group, the significance of the differences (F) and the percent of variance accounted for by the model (R^2) can then be calculated. These statistics coupled with the Monte Carlo Simulation, mentioned earlier, provide a metric of the successfulness of any resulting model (Wooldridge and Helms, 1976; Vreuls and Wooldridge, 1981).

Unit weighting (Wainer, 1976; Winter, 1974) is particularly useful when the number of measures under consideration is very large relative to the number of data points. It provides a conservative (non-optimal) discrimination which does not suffer from shrinkage as severely as parametric methods do, and the analysis does not consume any degrees of freedom. In some instances, where multivariate distribution assumptions are violated, this technique can out perform the elaborate parametric techniques.

Once a parametric discriminant function has been determined a non parametric function can be easily computed for the same variables and can be used as a benchmark for comparison.

Reliability Tests. Along with mathematical indicators of significance, the probability of misclassification can be estimated using a Monte Carlo simulation or a cross-validation of the system using actual performance data. The following paragraphs describe the procedures that can be used.

Mathematical Basis for Monte Carlo Simulation. As a first consideration, the simulated data must have a multivariate distribution much like the original sample. For the most part, the tax return variables in the analysis have already been transformed and are expected to be normally distributed. This mixture of distributions will be satisfactorily simulated with random vectors generated from a multivariate normal population. Since the interrelationships existing in the performance measures is the undesirable characteristic responsible for the development of 'ridge' analyses, it is crucial that the same variance-covariance structure of the data be represented in the simulation. Thus it is important not only to simulate a multivariate normal population, but also to be able to specify the variance-covariance matrix of the population.

Scheuer and Stoller (1960) suggest a method for generating random vectors from a multivariate normal population with a specified variance covariance matrix based on matrix equations. To simplify description of the technique, it will be assumed at first that the mean of the random vectors is zero. The result is no loss in generality, for a vector x with a mean of zero and a variance-covariance matrix Σ , the vector $x+\mu$ has the same variance-covariance matrix Σ and mean vector μ. It is then possible to concentrate on generating a random vector $x=(x_1, x_2, \ldots x_n)$ from $N(0,\Sigma)$, the multivariate normal distribution with mean vector zero and variance-covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma 11 & \ldots \sigma 1n \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \sigma n1 & \ldots \sigma nn \end{bmatrix}.$$

Let y be distributed $N(0,I_n)$, where I_n is the unit matrix of size n, and let $x=Cy$. Then x is distributed $N(0,CC')$. It is required that CC' be equal to Σ in this case. The matrix C is unique and readily determined if C is lower triangular. The elements of C are determined recursively as follows:

$$C_{i1} = \sigma_{i1} \Big/ \sqrt{\sigma_{11}}, \quad 1 \leqslant i \leqslant n,$$

$$C_{ij} = \sqrt{\sigma_{ii} - \sum_{k=1}^{i-1} C^2_{ik}}, \quad 1 \leqslant i \leqslant n,$$

$$C_{ij} = \left[\sigma_{ij} - \sum_{k=1}^{j-1} C_{ik} C_{ik} \right] \Big/ C_{jj}, \; 1 < j < i \leqslant n,$$

$$C_{ij} = 0, \qquad 1 < j \leqslant n.$$

This technique is referred to as the "square root" method and C is the "square root" of Σ.

Once C has been determined, x is obtained by

$$x_i = \sum_{j=1}^{i} C_{ij}y_j, \; i = 1, \ldots, n$$

where $y_1, \ldots y_n$ are independent standard normal variables, N(0, 1).

Box and Miller (1958) suggest a method for computation of random normal deviates. This approach has been shown to be more accurate than other known methods for generating normal deviates from independent random numbers: (1) the inverse Gaussian function of the uniform deviates, (2) Teichroew's approach, (3) a rational approximation such as that developed by Hastings, (4) the sum of a fixed number of uniform deviates, and (5) rejection type approach.

The method may be used to generate a pair of random deviates from the same normal distribution starting from a pair of random numbers. Letting U_1 and U_2 be independent variables from the same rectangular density function on the interval (0,1),

$$y_1 = (-2 \operatorname{Log}_e U_1)^{1/2} \operatorname{COS} 2\pi U_2 \text{ and}$$

$$y_2 = (-2 \operatorname{Log}_e U_1)^{1/2} \operatorname{SIN} 2\pi U_2$$

provides a pair of independent random variables, $(y_1, y2)$, from the same normal distribution with mean zero, and unit variance.

The new random vector x can now be computed given the original μ and Σ as

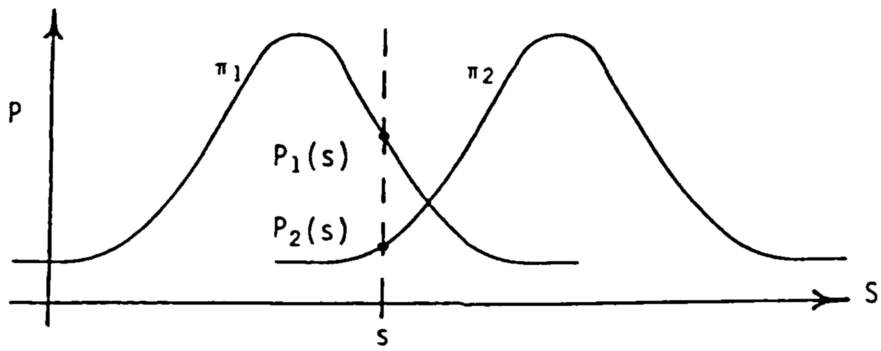
$$x_i = \left[\sum_{j=1}^{i} C_{ij}\, y_i \right] + \mu_i, \; i = 1, \ldots, n.$$

A simulation program computes a thousand new independent random vectors for each group with the same means and variance-covariance matrix as the actual returns and classifies them as to their respective populations using the chosen discriminant function. The resulting misclassifications can then be used to compare discriminant functions by calculating their respective percent error of classification. An acceptable criterion was also developed to optimally classify vectors for each discriminant function.

The score s will be calculated by:

$$s = b_1x_1 + b_2x_2 + b_3x_3 + \ldots + b_nx_n,$$

where the b's are the discriminant coefficients for n number of performance measures or x's. In other words, the b's make up the discriminant function for the sample vector composed of n number of x's. s is now a new variate in discriminant space belonging to one of two distributions in this case. The cost function can be derived using the score distributions in discriminant space of the sample used to empirically develop the discriminant function, s. There will be two of these distributions: $¶_1$ and $¶_2$. Any single score, s, will then fall somewhere in these distributions, with probabilities of $P_1(s)$ and $P_2(s)$.



In classifying s as belonging to distribution π_1 or π_2, two errors can be made. If s actually belonged to distribution π_1, an error would be made if s was classified as belong to π_2 (see Figure 5).

On the other hand, s could belong to π_2 and be classified as belonging to π_1. There is a cost associated with each type of error. Let C(2/1) be the cost of the first type of error and C(1/2) be the cost of the second. Table 2 is a logic table of the costs of correct and incorrect classification. It follows, that an effective classification scheme should minimize the cost of classification.

TABLE 2. CLASSIFICATION

		$\pi 1$	$\pi 2$
Actual	$\pi 1$	0	$C(2\|1)$
Membership	$\pi 2$	$C(1\|2)$	0

If we select a score, s, in discriminant space, the potential cost of using that point for classification can be estimated. The probability that s will be classified as belonging to π_2 even though it belongs to π_1 is

$$P(2 \mid 1,s) = \int_S^{\infty} P_1(s)ds.$$

Given that we already have an actual sample distribution where the number of observation in π_1, n_1, is known and the number of scores less than s in π_1 can be summed as m_1,

$$P(2 \mid 1,s) = \frac{n_1 - m_1}{n_1}.$$

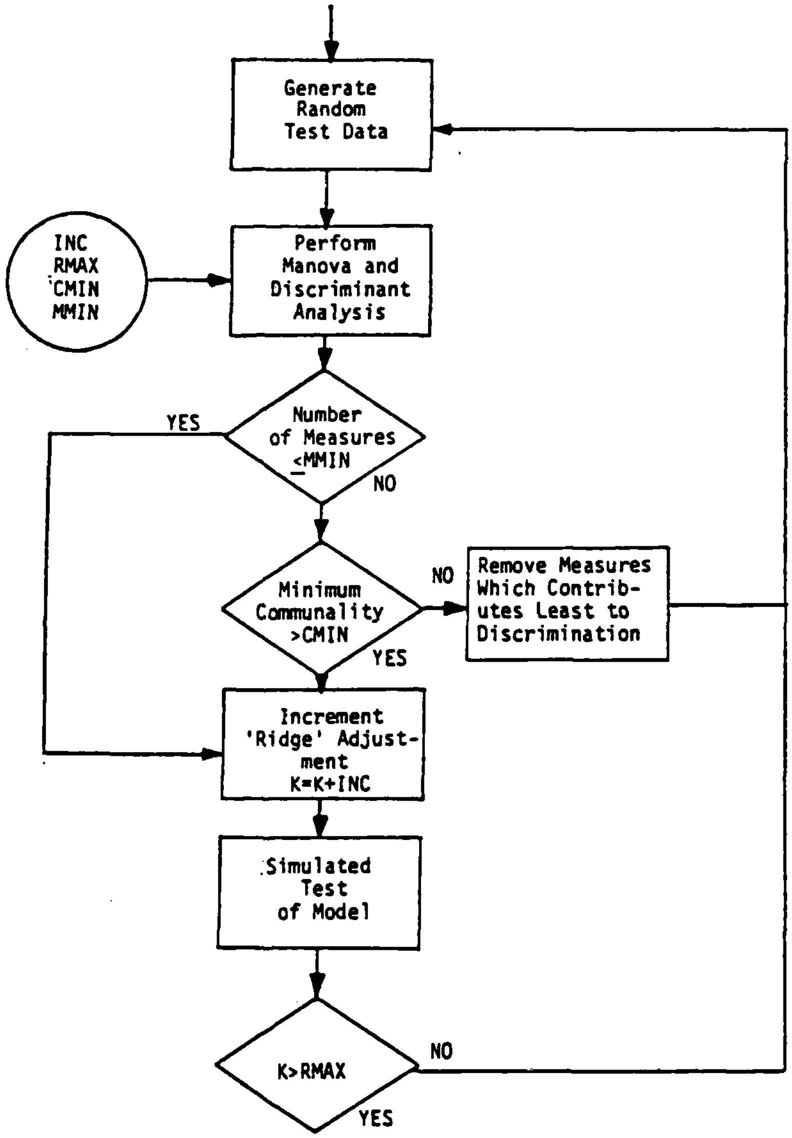


Figure 5. 'Ridge' Adjusted Discriminant Analysis Functional Flow.

The probability of misclassification of an observation from $\P_2$ is then

$$P\,(1 \mid 2,s) = \int_{-\infty}^{S} P_2(s)ds = \left(1 - \frac{n_2 - m_2}{n_2}\right),$$

where n_2 is the number of observations in $\P_2$ and m_2 is the number of scores in $\P_2$ less than s.

The probability of a $\P_1$ pilot achieving a particular score, s, is $P_1(s)$. This probability can be calculated using the number of observations in $\P_2$ falling in the period of integration bounded by s, $m_{s,1}$, divided by n_1 or simply

$$P_1(s) = \frac{m_{s,1}}{n_1}.$$

Thus, the probability associated with misclassifying a score from $\P_1$ is

$$P_1(s)P(2 \mid 1,s) \text{ or } \frac{m_{s,1}}{n_1}\;\frac{n_1 - m_1}{n_1}$$

and the probability of misclassifying a score from $\P_2$ is

$$P_2(s)P(1 \mid 2,s) \text{ or } \frac{m_{s,2}}{n_2}\left(1 - \frac{n_2 - m_2}{n_2}\right).$$

The average or expected loss from costs of misclassification is the sum-of-the-products costs of each misclassification multiplied by the probability of its occurrence;

$$C(2 \mid 1)P(2 \mid 1,s)P_1(s) + C(1 \mid 2)P(1 \mid 2,s)P_2(s), \text{ or}$$

$$C(2 \mid 1)\;\frac{n_1 - m_1}{n_1}\;\frac{m_{s,1}}{n_1} + C(1 \mid 2)\left(1 - \frac{n_2 - m_2}{n_2}\right)\frac{m_{s,2}}{n_2}$$

If $C(1 \mid 2) = C(2 \mid 1) = 1$, the expected loss is

$$\frac{n_1 - m_1}{n_1}\;\frac{m_{1,s}}{n_1} + \left(1 - \frac{n_2 - m_2}{n_2}\right)\frac{m_{s,2}}{n_2}$$

Assuming $\P_1$ and $\P_2$ are normally distributed, for a given score s, the probability of misclassification is minimized by assigning s to the sample that has the higher conditional probability. Thus, the rule is:

$$\P_1 \text{ is chosen if } \quad \frac{n_1 - m_1}{n_1} \quad \frac{m_{s,1}}{n_1} \geqslant \left(\frac{1 - n_2 - m_2}{n_2} \right) \frac{m_{s,2}}{n_2}$$

$$\text{and } \P_2 \text{ is chosen if } \quad \frac{n_1 - m_1}{n_1} \quad \frac{m_{s,1}}{n_1} < \left(\frac{1 - n_2 - m_2}{n_2} \right) \frac{m_{s,2}}{n_2}$$

This line of reasoning follows closely the derivation put forth by Anderson (1958) for discriminant classification criteria. For out purposes, it suffers from two fatal problems. First, the computation of $P_1(s)$ and $P_2(s)$ relies on the arbitrary determination of a period of integration. Since these values can change directly with the length of the integration period selected, the cost function itself becomes a problem to define. Secondly, the distributions to be analyzed are known not to be normally distributed. This upsets the decision rule described above.

A slight departure from these results provides a practical decision algorithm for realistic data analysis conditions. Assuming that the intrinsic cost associated with misclassificaitons of both kinds is still 1, the total probability of misclassification for any s provides a relative metric for comparison. The total probabiltiy of misclassification is

$$P(1 \mid 2,s) + P(2 \mid 1,s) \text{ or } \frac{n_1 - m_1}{n_1} + \left(\frac{1 - n_2 - m_2}{n_2} \right)$$

For an existing sample, P(1 2,s) and P(2 1,s) are simply the percentage of misclassifications of $\P_1$ and $\P_2$ respectively. Their sum is the total percentage of misclassifications for s in discriminant space. Given a sample, this total percentage can be evaluated along the entire discriminant space to locate the minimum error that will occur somewhere between the means of $\P_1$ and $\P_2$. The value in discriminant space where the minimum error occurs will be called the break even point.

Several potential discriminant functions can be derived for each sample of data. A classification criteria for each function can be determined by finding the break even point using the sample discriminant scores, and the discriminant functions can then be compared using their respective percentage of classification errors on generated independent variates resembling the original sample. This simulation can lend itself to the selection of K in a 'ridge' discriminant analysis as well as to the estimation of the power of the final discriminant function. Figure 6 shows the functional flow for the Monte Carlo procedure.

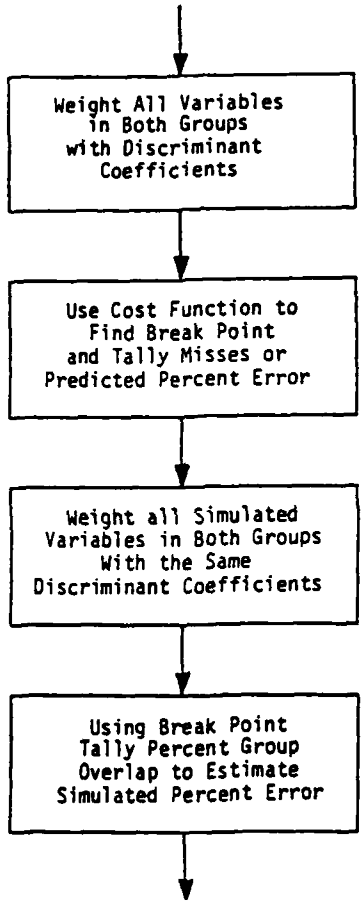


Figure 6. Simulated Test of Discriminant Model Functional Flow.

Cross-Validation of Discriminant Function. There are some extremely sophisticated, iterative cross-validation procedures suggested by the literature (Mosteller, Tukey, 1977). Considering this limited number of observations usually available in most development studies, the efficacy of many of these methods would have to be questioned. A double crossvalidation procedure is advisable when the sample size is sufficient. This procedure will require that 10 percent of the data from each group be set aside before any statistical analyses or processing and used only for testing purposes. This will assure that the return data is different or independent from those used to guide the choice of the functions form and from those used to choose its numerical coefficients. The reserved samples will be inserted into the discriminant system and the results will be computed much the same as the simulation test.

STATISTICAL ISSUES IN THE ASSESSMENT OF MULTIVARIATE MODELS

How much data constitutes a sufficient sample of performance? This question continually haunts most researchers. In fact, scientists are often classified as to their preferred criterion of sample size. On one hand you have those that pursue the 'holy grail' of statistical significance at any cost. On the other hand, there are those who shun conventional hypothesis testing and prefer to describe the complete environment with holistic response surface techniques; requiring many variables and few observations. It is important to remember for the purposes of this discussion that we are not so interested in the assessment of performance as we are in gauging the relative effectiveness of measures and measurement systems. The practical significance of the improvement between an existing measurement system and any new system in describing measurement space is the final judgment to be made. This section is confined to the intelligent estimation of the confidence with which a resulting measurement system can be applied rather than, purely, the achievement of statisitical significance.

Power and Significance. The validity of a measurement model cannot be assessed by statistics, but depends on the adequacy of sampling, and the quality of the data collection and analysis (Simon, 1981). Tests of significance, when properly employed, estimate the probability of mistakenly rejecting the null hypothesis, or Type I error. It is not unusual that a researcher accepts a hypothesis based solely on indications of an F-test or t-test. It has also been a practice to force statistical significance by increasing the number of samples, by changing the parameters of the test, or selecting another form of test.

Unfortunately, even under the best of conditions tests of significance provide little interpretive information. Inferential statistics tend to concentrate on Type I error and ignore the risk of committing Type II errors. In other words, experimenters using significance tests often do not weigh the risk of saying there is no effect when, in fact, there is.

When designing measurement systems it is important to weigh both risks in the interpretation of effects. The need for additional data may be the most important determination. This is where the concept of statistical

power comes in. The power of an inference is computed as 1-B, where B is the probability of committing a Type II error. Textbooks, conventionally, describe ways to compute the power of the F-test and plot the operating characteristics based on sample size. Selecting a sample size is then merely a problem of weighing the cost of increasing the number of observations against the costs of the Type I and Type II errors expected. The underlying problem with this seemingly simple concept is that the variance of the groups must be known in advance of the study. At the onset of a measurement development study the behaviour of the variables and measurement systems in their particular operational setting are not well defined. The study design must, therefore, be adaptive to allow estimation of the variability of the performance measurement systems, the resulting power of the statistical inferences derived, and the cost effectiveness practicality of collecting more data.

In the back of this appendix contains subroutines for calculating the probability and the power of the F-test. When computing the power of the F-test (Woodward and Overall, 1976) it is important to determine the noncentrality parameter

$$\lambda = \frac{n \Sigma T_j^2}{\sigma_E^2}$$

(Winer, 1971) where $X_{ij} = u + T_j + E_{ij}$.

under the hypothesis that

$T_1 = T_2 = \ldots = T_k = 0$, it follows that = 0 and the distribution of the F-ration has the central F distribution. When = 0 the distribution is noncentral and depends upon the parameter .

Another comparable method for computing power is to employ the multivariate effect size: Suggested by Shaffer and Gillo (1974):

$$ES = \sqrt{\frac{Cr^2}{1-Cr^2}}$$

using the correlation ratio, Cr. The discriminant analysis program developed by Cooledy and Lohnes (1971) calculates the correlation ratio:

$$Rj2 = \frac{\lambda j}{1 + \lambda j}$$

where j is the eigen value of the jth discriminant function. The correlation ratio suggested by Shaffer and Gillo is

$$Cr = \frac{Tr\ (BW^{-1})}{p+Tr(BW^{-1})}$$

where $Tr(BW^{-1})$ is the trace of the product of the within and between groups sums of squares and cross products and

p is the number of roots. Cr is equivalent to the R^2 above for the two-group case. This fact allows us to side step, for the case of measurement system development, discussion of the controversy over the calculation of R^2.

Nevertheless, a desirable sample size can be determined from the resulting ES using tables developed by Cohen (1971).

Shrinkage. In review, the correlation ratio, R^2, estimates the proportion of variance explained by the predictor variance. The correlation ratio is an extremely useful interpretive tool, but it is also affected by sample size. If the number of predictors/measures M is appreciable, relative to the number of observations N, the sample value of R^2 is biased upward. Several adjustments are suggested in the literature (Lane, 1971)

$$\hat{R}^2_{\text{Wherry}} = 1 - (1-R^2)\left(\frac{N-1}{N-M-1}\right)$$

$$\hat{R}^2_{\text{N-Lord}} = 1 - (1-R^2)\left(\frac{N-1}{N-M-1}\right)\left(\frac{N+M-1}{N}\right)$$

$$\hat{R}^2_{\text{Darlington}} = 1 - (1-R^2)\left(\frac{N-1}{N-M-1}\right)\left(\frac{N+1}{N}\right)\left(\frac{N-2}{N-M-2}\right)$$

It is suggested by Cohen and Cohen (1975) that particularly for stepwise regression the adjustment for R^2 should be:

$$\hat{R}^2 = 1 - (1-R^2)\left(\frac{N-1}{N-k-1}\right)$$

where k is the total number of measures from which the selection was made. In the case of stepwise regression, where measures are chosen one at a t e on the basis of their relationships with Y, R^2 will tend to be too large because of the tendency of the stepwise procedure to capitalize on chance.

Research using the various correction formulas has shown that as the ratio of measures to observations increases the adjusted R^2 becomes less meaningful. There is no single solution that satisfies all conditions. Stepwise procedures are particularly susceptible to error as the large number of initial variables are not amenable to conventional multivariate selection techniques. As the number of measures is severely reduced, the shrinkage of R^2 is directly related to the M/N ratio and all the suggested corrections perform equally satisfactorily. The simulator techniques discussed earlier can be used in conjunction with the adjusted correlation ratio to judge the effectiveness of the measurement system.

Repeated Measures. Conditions in the operational environment often cause the empirical researcher to violate traditional experimental design

standards. In training and performance measurement studies it is often necessary to collect data on the said trainees both before and after training. Repeating measurement of each trainee for each level of training may also be required when the population of trainees is sparse in relation to the total number of candidate performance measures. These data collection procedures violate both the within groups and between groups assumptions of independent observations.

Various methods for adjusting scores for repeated measures have been proposed, although Weinman (Wooldridge, Breaux and Weinman, 1976) developed a theoretical model that consumes the least degrees of freedom. The rest of this section will be devoted to discussion of Weinman's conservative model.

Consider that each trainee/subject was observed repeatedly in each training condition or group for every group. Let X_{mijk} be the score on measure m, in group i, subject j and trial k. A model for such a score can be stated as

$$X_{mijk} = \mu m + \alpha i + \beta j + (\alpha \beta)ij + k(i) + \beta\alpha jk(i)$$

where:

μm = grand mean of measure over groups, trials and subjects,

αi = group effect,

βj = subject effect,

$\alpha k(i)$ = effect of trial within group i,

$(\alpha\beta)_{ij}$ = interaction effect for subject j and group i,

$(\beta\alpha)_{jk(i)}$ = interaction effect for subject j and trial k on group i.

Estimates for these parameters and the degrees of freedom associated with each are:

$$\mu m = \overline{X}_{m}\cdots$$

$$\alpha i = \overline{X}_{mi}\cdot\cdot - \overline{X}_{m}\cdots$$

$$\beta j = \overline{X}_{m}\cdot j\cdot - \overline{X}_{m}\cdots$$

$$\alpha k(i) = \overline{X}\cdot i\cdot k - \overline{X}_{mi}\cdot\cdot$$

$$(\alpha\ \beta)_{ij} = \overline{X}_{mij}\cdot - \overline{X}_{m}\cdot j\cdot - \overline{X}_{mi}\cdot\cdot + \overline{X}_{m}\cdots$$

$$(\beta\ \alpha)j^{k}(i) = \overline{X}_{mijk} - \overline{X}_{mij}\cdot - \overline{X}_{mi}\cdot k + \overline{X}_{mi}\cdot\cdot$$

where the dot (.) replacing a subscript indicates the subscript being summed. To keep sufficient degrees of freedom, it is necessary to assume

the interaction effect for subject, trial and day, $(\ )_{jk(i)}$, to be always zero. The effects ji, k(i), and $(\ )_{ij}$ can be estimated and removed from the scores X_{mijk}.

The program in the appendix replaces each score with X'mijk where

$$X'mijk = Xmijk - \overline{X}mi.k - \overline{X}mij. + 2\overline{X}mi..$$

The new scores can be used in further multivariate analyses.

There are a few drawbacks to this approach. First, the model is very sensitive to outliers or instability in the data, due to the requisite assumption of the nonexistent three way interaction on no error term. Secondly, the power of the resulting measurement system is somewhat difficult to estimate, as the actual degrees of freedom are unknown after the removal of portions of subject variance. The degrees of freedom could range anywhere from one less than the actual number of subjects (in which case little power has been gained by the repeated measures design) to one less than the total number of observations per group (assuming groups of equal size).

REFERENCES

Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: John Wiley & Sons, 1958.

Bittner, A. C., Jr. Reduction in Potential User Population as the Result of Imposed Anthropometric Limits: Monte Carlo Estimation. TP-74-6. Point Mugu, California: Naval Missile Test Center, 1974.

Box, G. E. P. and Miller, M. E. "A Note on Generation of Random Normal Deviates." Annals of Mathematical Statistics. 29, 1958: 610.

Bruning, J. and Kintx, B. Computational Handbook of Statistics. Illinois: Scott, Foresman & Company, 1977.

Cohen, J., and Cohen, P. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1975.

Cohen, Jacob. Statistical Power Analysis for the Behavioral Science. Academic Press, New York, 1971.

Cooley, W. W. and Lohnes, P. R. Multivariate Data Analysis. New York: John Wiley & Sons, 1971.

Dixon, W. and Massey, F. J., Jr. Introduction to Statistical Analysis. New York: McGraw-Hill Book Company, 1969.

Farebrother, R. W. "The Minimum Mean-Square-Error Linear Estimator and Ridge Regression." Technometrics. 17, 1975: 127-128.

Gibra, I. N. Probability and Statistical Inference for Scientists and Engineers. Prentice-Hall, Englewood Cliffs, N.J., 1973.

GNANADESIKAN, R. and Wilk, M. B. "Data Analytic Methods in Multivariate Statistical Analysis". Anals of Mathematical Statistics, 1970.

Gordon, G. System Simulation. Englewood Cliffs, N. J.: Prentice Hall, 1969.

Hays, W. Statistics for the Social Sciences. New York: Holt, Rinehart and Winston, Inc., 1973.

Hoerl, A. E. and Kennard, R. W. "Ridge Regression: Biased Estimation for Nonorthogonal Problems." Technometrics. 12, 1970: 55-67.

Hoerl, A. E. and Kennard, R. W. "Ridge Regression: Applications to Nonorthogonal Problems." Technometrics. 12, 1970: 69-82.

Kelly, M. J., Wooldridge, A. L., Hennessy, R. T., Vreuls, D., Barneby, S. F., Cotton, J. C., and Reed, J. C. Air Combat Maneuvering Performance Measurement. AFHRL-TR-79-3. Williams AFB, AZ: Flying Training Division, Air Force Human Resources Laboratory, September 1979.

Lane, N. E. The Influence of Selected Factors on Shrinkage and Overfit in Multiple Correlation. Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1971.

Neave, H. R. "Quick and Simple Test Based on Extreme Observations." Journal of Quality Technology. Volume 11, No. 2, April 1979: 66-79.

Scheuer, E. M. and Stoller, D. S. "On the Generation of Normal Random Vectors. Journal of the American Statistical Association. 1960: 278-281.

Simon, C. W. Methods for Improving Information from 'Undesigned' Human Factors Experiments. TRP75-287. Culver City, CA: Hughes Aircraft Company, 1975.

Simon, C. W. Applications of Advanced Experimental Methods To Visual Technology Research Simulator Studies: Supplemental Techniques. NAVTRAEQUIPCEN 78-C-0060-3, Naval Training Equipment Center, Orlando, Florida, 1981.

Tatsuoka, M. M. Discriminant Analysis-The Study of Group Differences. Champaign, IL: The Institute for Personality and Ability Testing, 1970.

Thomas, D. G. Computer Methods for Generating Pseudo-Random Numbers from Pearson Distributions. Blacksburg, VA: Virginia Polytechnic, May 1966.

Vreuls, D., Wooldridge, L., Obermayer, R., Johnson, R., Norman, D., and Goldstein, I. Development and Evaluation of Trainee Performance Measures in an Automated Instrument Flight Maneuvers Trainer. NAVTRAEQUIPCEN 74-C-0063-1. Orlando, FL: U. S. Naval Training Equipment Center, October 1976.

Wainer, H. "Estimating Coefficients in Linear Models: It Don't Make No Nevermind." Psychological Bulletin. Volume 83, 1976: 213-217.

Winer, B. J. Statistical Principles In Experimental Design. McGraw-Hill, New York, 1971.

Winter, B. B. "Nonparametric Density Estimation and Statistical Discrimination." Psychological Bulletin. Volume 81, No. 6, 1974: 371-379.

Woodward, J. A., Overall, J. E. A Computer Program for Calculating Power of the F-Test. Educational and Psychological Measurement, 1976, 165-168.

Wooldridge, L. A Monte-Carlo Approach to Ridge Discriminant Analysis. Master's Thesis, Orlando, FL: University of Central Florida, 1980.

Wooldridge, A. L., Breaux, R., Weinman, D. G. Statistical Issues in the Use of Multivariate Methods for the Selection of Flight Simulator Performance Measures. NAVTRAEQUIPCEN 75-C-0091-1, 1976.

SUBROUTINES

PROGRAM C-1. TUKEY QUICK TEST OF LOCATION AND T-TEST.

This disc-bound version of the TUKEY test was written for an extremely large number of initial measures (up to 1000).

```
 SUBROUTINE SSATUK (OPTN,SIGLVL)
C
C  The TUKEY test is a nonparametric (or distribution free) statistical
C  test. This program computes the Tukey test or the student t-Test for
C  two groups of data and the associated significance levels.
C
 INCLUDE 'SSA.DEF/NOLIST'
C
 REAL MIN1,MAX1,MIN2,MAX2,MINP,MAXP
 REAL YMEAN(2),YSD(2),YSUM(2),YSQSM(2)
 INTEGER TUK,OPTN,SIGLVL
 LOGICAL*1 GEORGE
 COMMON/POOL/ MINP,MAXP,MING,MAXG
C
 CLOSE(UNIT=TEMP1C,DISP='SAVE')
 OPEN(UNIT=TEMP1C,NAME='TEMP1.SCR',ACCESS='DIRECT',ASSOCIATEVARIABLE
 1=KK,TYPE='OLD',FORM='UNFORMATTED',RECORDSIZE=24)
C
C  Determination of the significant measures.
C
 KK = 1   ! Reposition Scratch file to Grp 1.
 KNT1 = 0
 KNT2 = 0
 KKK = KK
 LVLS = 1000  ! A standard to compare to.
 REWIND WORKC
 READ (WORKC) NUMG
 READ (WORKC) NMEAS
 JGRP2 = 2*NMEAS
 IF (OPTN .GT. 4) WRITE (LPOUTC,1000)
1000 FORMAT('1 "t-Test" results : '/' Measure',12X,'t',
 16X,'Significant?')
   DO 50 I = 1,NMEAS
   READ(TEMP1C'KK) XBAR1,SD,NMOBS1,PSUM,RANG,VAR1
   READ(TEMP1C'KK) KNT1,MIN1,MAX1,MINP,MAXP
D WRITE(5,957) XBAR1,SD,NMOBS1,PSUM,RANG,VAR1
D WRITE(5,937) KNT1,MIN1,MAX1,MINP,MAXP
937 FORMAT(1X,I4,4(G13.6,1X))
   KK = JGRP2 + KKK
   READ(TEMP1C'KK) XBAR2,SD,NMOBS2,PSUM,RANG,VAR2
   READ(TEMP1C'KK) KNT2,MIN2,MAX2,MING,MAXG ! Get the Grp2 data.
D WRITE(5,957) XBAR2,SD,NMOBS2,PSUM,RANG,VAR2
D WRITE(5,917) KNT2,MIN2,MAX2,MING,MAXG
917 FORMAT(1X,I4,2(1X,G13.6),2(1X,I6))
C
C  Calculate the Degrees of Freedom for both Grps and for pooled.
C
   X1 = NMOBS1
```

```
   DF1 = NUMG - 1
   X2 = NMOBS2
   DF = X1 + X2 - 2.0
   IDF = DF
   REC = (1.0/X1) + (1.0/X2)
C
C  Compute the "t" level of significance and save.
C
   DENOMT = SQRT ((VAR1/X1) + (VAR2/X2))
   IF (DENOMT .EQ. 0) GO TO 10 !Prevents "Floating Zero Divide" error.
   GO TO 20
10   T = 1.0
   GO TO 30
20   T = ABS((XBAR1-XBAR2) / (DENOMT))
30   KK = JGRP2 + KKK
   TUK = 0
   WRITE(TEMP1C'KK) XBAR2,SD,TUK,T,RANG,VAR2 ! Save the "t" stat
D WRITE(5,957) XBAR2,SD,TUK,T,RANG,VAR2
957 FORMAT(1X,2(1X,F10.5),1X,I5,3(1X,G13.6))
   IF(OPTN .LT. 5) GO TO 40    ! Loop for TUKEY.
   GEORGE = TSTSIG((T*2.0),IDF,SIGLVL)    ! George is a logical.
   IF (GEORGE) TUK = 1000
   KK = JGRP2 + KKK
   WRITE(TEMP1C'KK) XBAR2,SD,TUK,T,RANG,VAR2 ! Save in Grp2.
   WRITE (LPOUTC,1100) I,T,GEORGE
1100   FORMAT(1X,I7,2X,G13.6,6X,L6)
D WRITE(5,957) XBAR2,SD,TUK,T,RANG,VAR2
40   KK = 2*I + 1
   KKK = KK
50   CONTINUE
C
 TTST = .05
 IF (OPTN .EQ. 2) TTST = .01
 IF (OPTN .EQ. 3) TTST = .005
C
C  The relative position of one group and the other should be the
C  same for all measures. This is determined by checking the value
C  of the Means and correcting as necessary .
C
 KK = 1
 KKK = KK
D WRITE(5,999)
999 FORMAT('0  SSATUK  --  Entering DO 80 ')
   DO 80 J = 1,NMEAS
   READ(TEMP1C'KK) XBAR1,SD,NMOBS,PSUM,RANG,VAR1
   READ(TEMP1C'KK) KNT1,MIN1,MAX1,MINP,MAXP
   KK = JGRP2 + KKK
   READ(TEMP1C'KK) XBAR2,SD,TUK,T,RANG,VAR2
   READ(TEMP1C'KK) KNT2,MIN2,MAX2,MING,MAXG
   IF (XBAR1 .LT. XBAR2) GO TO 60
   GO TO 70
C
C  The negative sign on the integer count variable will be used
C  to adjust for differences in position.
C
```

```
60   KNT2 = -KNT2
70   KK = JGRP2 + KKK
   WRITE(TEMP1C'KK) XBAR2,SD,TUK,T,RANG,VAR2
   WRITE(TEMP1C'KK) KNT2,MIN2,MAX2,MING,MAXG
   KK = 2*J + 1
   KKK = KK
80   CONTINUE
C
 MESUSD = 1  ! The count of the measures used.
 KK = JGRP2 + 1
D WRITE(5,997)
997 FORMAT('0 SSATUK  --  Entering DO 90  ')
   DO 90 J = 1,NMEAS
   READ(TEMP1C'KK) XBAR2,SD,TUK,T,RANG,VAR2
   READ(TEMP1C'KK) KNT2,MIN2,MAX2,MINP,MAXP
   IF (TUK .GE. LVLS) MESUSD = MESUSD + 1
90   CONTINUE
D WRITE(5,981) MESUSD
981 FORMAT('0   MESUSD = ',I5)
 IF (OPTN .GT. 4) GO TO 230  ! Use "t" significance.
C
C  Since this is the TUKEY Procedure, we have to count the # of X's
C  from the Grp having the lower Mean that are below the minimum X in the
C  Grp having the higher Mean. Then we count those X's from the Grp
C  having the higher Mean that are above the Max X of the Grp having the
C  lower Mean. Confused?
C
C   The probability equation is:
C
C
C        2 * XTMP   ( XTMP ** TUK-1 )
C P(TUK) =     ------------  *  --------------------
C                     XTMP**2 - 1      (XTMP + 1) ** TUK
C
C  where XTMP is the # of observations in Grp2 divided by the #
C  of observations in Grp1. (NOTE: The higher "N" is always the
C  numerator.) In the case of equal N's, XTMP = (N+1)/N.
C
C      & TUK is the # of significant differences for a given meas.
C
C
 XTMP = X2 / X1
 IF (XTMP .LT. 1.0) XTMP = X1 / X2
 IF (X1 .EQ. X2) XTMP = (X1+1.0) / X2
 XMULT = (2.0 * XTMP) / (XTMP ** 2 - 1.)
C
 REWIND WORKC
 READ (WORKC) NUMG
 READ (WORKC) NMEAS
 KKK = 1
 MKK = 1
 WRITE (LPOUTC,1200)
1200 FORMAT('1',80('*')/10('*'),10X,'  TUKEY QUICK TEST OF LOCATION
 1 SUMMARY ',10X,10('*')/80('*'))
 WRITE (LPOUTC,1300)
```

```
1300 FORMAT('0Measure * Pooled Minimum',5X,'Range',5X,'* TUKEY
 1 Count * Level of Significance')
      DO 200 I = 1,NMEAS
      KK = KKK
      READ (TEMP1C'KK) XBAR1,SD,NMOBS1,PSUM,RANG,VAR1
      READ(TEMP1C'KK) KNT1,MIN1,MAX1,MINP,MAXP
D WRITE(5,957) XBAR1,SD,NMOBS1,PSUM,RANG,VAR1
D WRITE(5,937) KNT1,MIN1,MAX1,MINP,MAXP
      KK = JGRP2 + KKK  !Get to begining Grp2.
      READ(TEMP1C'KK) XBAR2,SD,TUK,T,RANG,VAR2
      READ(TEMP1C'KK) KNT2,MIN2,MAX2,MING,MAXG
D WRITE(5,957) XBAR2,SD,TUK,T,RANG,VAR2
D WRITE(5,917) KNT2,MIN2,MAX2,MING,MAXG
      IF (MING .EQ. MAXG) GO TO 190 ! TUKEY doesn't apply.
      IF (MING .EQ. 1) GO TO 100 !Count # of meas.
      GO TO 130
C
100      READ(WORKC) NUMOBS
      TUK = 0
      DO 110 K = 1,NUMOBS
      READ(WORKC) (X(NN) , NN = 1,NMEAS)
      IF (X(I) .LT. MIN2) TUK = TUK + 1
110      CONTINUE
C
      READ (WORKC) NUMOBS
      DO 120 K = 1,NUMOBS
      READ (WORKC) (X(NN), NN = 1,NMEAS)
      IF (X(I) .GT. MAX1) TUK = TUK + 1
120      CONTINUE
C
      GO TO 160
130      READ(WORKC) NUMOBS
      TUK = 0
C
      DO 140 K = 1,NUMOBS
      READ(WORKC) (X(NN) , NN = 1,NMEAS)
      IF (X(I) .GT. MAX2) TUK = TUK + 1
140      CONTINUE
C
      READ (WORKC) NUMOBS
      DO 150 K = 1,NUMOBS
      READ (WORKC) ( X(NN),NN = 1,NMEAS)
      IF (X(I) .LT. MIN1) TUK = TUK + 1
150      CONTINUE
C
C  Now calculate the Level of Significance value.
C
160      REWIND WORKC
      READ (WORKC) NUMG
      READ (WORKC) NMEAS
          SVAL = 1.0
      IF (TUK .LT. 20) GO TO 170     !Prevents "Floating Overflow" error.
      GO TO 180
170      IF (TUK .NE. 0) SVAL = XMULT * ((XTMP**TUK-1)/
 1 (( XTMP + 1.0) ** TUK))
```

```
180      WRITE(LPOUTC,1400) I,MINP,RANG,TUK,SVAL
1400      FORMAT(1X,I7,2(2X,G13.6),6X,I6,4X,F8.4)
     GO TO 195
190      SVAL = 1.0
195      KK = MKK
     READ (TEMP1C'KK) XBAR1,SD1,NMOBS1,PSUM,SQDSM,VAR
     KK = MKK
     WRITE (TEMP1C'KK) XBAR1,SD1,NMOBS1,PSUM,SVAL,VAR
     KKK = 2*I + 1
     MKK = KKK
200      CONTINUE
C
 WRITE (TTOUTC,1500)
1500 FORMAT('0Indicate the level of significance desired to keep
 1those measures.')
 READ (TTINC,1600) TTST
1600 FORMAT(F6.4)
 KKK = 1
C
   DO 220 I = 1,NMEAS
   KK = KKK
   READ (TEMP1C'KK) XBAR1,SD1,NMOBS1,PSUM,SVAL,VAR
   READ (TEMP1C'KK)
   IF (SVAL .GT. TTST) GO TO 210  !Prob. to high.
   MESUSD = MESUSD + 1
   KK = JGRP2 + KKK
   READ (TEMP1C'KK) XBAR2,SD2,TUK,T,RANG,VAR
   KK = JGRP2 + KKK
   TUK = 1000
   WRITE (TEMP1C'KK) XBAR2,SD2,TUK,T,RANG,VAR
210   KKK = 2*I + 1
220   CONTINUE
C
D KK=1
D NEND = NMEAS
D WRITE(5,975)
D975 FORMAT('  SSATUK --- ENTERING THE DO 979/977 BLOCK.')
D DO 979 L=1,NEND
D READ(TEMP1C'KK)XBAR,SDEV,NMOBS,PSUM,SVAL,VAR
D READ(TEMP1C'KK)KNT1,MIN1,MAX1,MINP,MAXP
D WRITE(5,957) XBAR,SDEV,NMOBS,PSUM,SVAL,VAR
D WRITE(5,937) KNT1,MIN1,MAX1,MINP,MAXP
D979 CONTINUE
C
D DO 977 L=1,NEND
D READ(TEMP1C'KK) XBAR2,SD2,TUK,T,RANG,VAR
D WRITE(5,957) XBAR2,SD2,TUK,T,RANG,VAR
D READ(TEMP1C'KK) KNT2,MIN2,MAX2,MING,MAXG
D WRITE(5,917) KNT2,MIN2,MAX2,MING,MAXG
D977 CONTINUE
C
C  Using the preceeding data we're now going to generate
C  the new data point, Y.
C
230 IF (MESUSD .GT. 1) GO TO 240
```

```
 WRITE(TTOUTC,1700)
1700 FORMAT('OSSATUK - No significant differences were found.')
 RETURN
C
240 IF (OPTN .GT. 4) GO TO 250  ! If "t-Test".
 GO TO 260
250 WRITE(LPOUTC,1800)
 WRITE(TTOUTC,1800)
1800 FORMAT('OThe new data point, Y, will be generated using'
 1/' the results of the "t-Test" ')
 GO TO 270
260 WRITE(LPOUTC,1900)
 WRITE(TTOUTC,1900)
1900 FORMAT('OThe new data point, Y, will be generated using'
 1/' the results of the TUKEY Quick Test of Location')
270 IF (MOD(OPTN,4) .EQ. 1) WRITE(LPOUTC,2000)
2000 FORMAT(' and Univariate SCALING.')
 IF (MOD(OPTN,4) .EQ. 2) WRITE(LPOUTC,2100)
2100 FORMAT(' and Univariate WEIGHTING.')
 IF (MOD(OPTN,4) .EQ. 3) WRITE(LPOUTC,2200)
2200 FORMAT(' and Translated SCALING.')
 IF (MOD(OPTN,4) .EQ. 0) WRITE(LPOUTC,2300)
2300 FORMAT(' and Translated WEIGHTING.')
 WRITE (LPOUTC,2400) TTST
2400 FORMAT('0 A significance level of ',F7.4,' was chosen.')
 KKK = 1
D WRITE(5,983)
983 FORMAT('0  SSATUK  --  Entering DO 280  ')
C
   DO 280 I = 1,NMEAS
   KK = KKK
   READ (TEMP1C'KK)
   READ (TEMP1C'KK) KNT1,MIN1,MAX1,MINP,MAXP
   IF ((OPTN.EQ.1) .OR. (OPTN.EQ.2) .OR. (OPTN.EQ.5) .OR.
  1  (OPTN.EQ.6)) MINP = 0.0
   KK = KKK
   READ (TEMP1C'KK)
   WRITE (TEMP1C'KK) KNT1,MIN1,MAX1,MINP,MAXP
   KK = JGRP2 + KKK
   READ(TEMP1C'KK) XBAR2,SD,TUK,T,RANG,VAR2
   IF (MOD(OPTN,2) .EQ. 0) RANG = SD ! We're weighting.
   KK = JGRP2 + KKK  !Get odd #'ed Grp2.
   WRITE (TEMP1C'KK) XBAR2,SD,TUK,T,RANG,VAR2
   KKK = 2*I + 1   ! Step thru Grp2 at odd #'s.
280   CONTINUE
C
 REWIND WORKC
 READ (WORKC) NUMG
 READ (WORKC) NMEAS
 TOBS = 0.0
C
D WRITE(5,989)
989 FORMAT('0  SSATUK  --  Entering DO 320.')
C
 DO 320 I = 1,2
```

```
 YSUM(I) = 0.0
 YSQSM(I) = 0.0
 READ(WORKC) NUMOBS
C
D WRITE(5,980)
980 FORMAT('   SSATUK  --  Entering DO 310 ')
 KKK = 1
    DO 310 J = 1,NUMOBS
   KK = KKK
   READ(WORKC) (X(NN),NN = 1,NMEAS)
   Y = 0.0
C
     DO 300 K = 1,NMEAS
     READ (TEMP1C'KK)
     READ (TEMP1C'KK) KNT1,MIN1,MAX1,MINP,MAXP
     KK = JGRP2 + KKK
     READ(TEMP1C'KK) XBAR2,SD2,TUK,T,RANG,VAR2
     READ(TEMP1C'KK) KNT2,MIN2,MAX2,MING,MAXG
     IF (TUK .LT. LVLS) GO TO 290
     XX = X(K)
     SIGN = 1.0
     IF (KNT2 .LT. 0) SIGN = -1.0
     IF (RANG .EQ. 0) GO TO 290
     Y = Y + (SIGN*(XX-MINP) / RANG)
290     KKK = 2*K + 1
     KK = KKK
300     CONTINUE
C
   IF ((I.EQ.1) .AND. (J.EQ.1)) WRITE (LPOUTC,2500)
2500   FORMAT('1   Measures contributing to the
 1 calculation of variable "Y"'/28X,'   GROUP      SUBJECT
 2       Y'/80('-'))
   WRITE (LPOUTC,2700) I,J,Y
2700   FORMAT(30X,I4,4X,I6,6X,G13.6)
   YSUM(I) = YSUM(I) + Y
   YSQSM(I) = YSQSM(I) + Y*Y
   KKK = 1
310   CONTINUE
C
 XOBS = NUMOBS
 TOBS = TOBS + XOBS
 YMEAN(I) = YSUM(I) / XOBS
 YSD(I) = (YSQSM(I) - ((YSUM(I)*YSUM(I)) / XOBS)) / (XOBS-1.)
320 CONTINUE
C
C  Calculate the pooled Y statistics.
C
D WRITE(5,987)
987 FORMAT('0 SSATUK - Calculating  the Y statistics.')
 YSUMP = YSUM(1) + YSUM(2)
 YSQSMP = YSQSM(1) + YSQSM(2)
 DF2 = TOBS - DF1
 YMEANP = YSUMP / TOBS
 YSDP = (YSQSMP - ((YSUMP*YSUMP) / TOBS)) / (TOBS-1.0)
 SST = YSQSMP - ((YSUMP*YSUMP) / TOBS)
```

```
      SSB = (X1*(YMEAN(1) - YMEANP)**2 + (X2*(YMEAN(2) - YMEANP)**2))
      SSW = SST - SSB
      F1 = (SSB/DF1) / (SSW/DF2)
      F2 = ((X2*(YMEAN(1) - YMEAN(2)))**2) / (YSD(1) + YSD(2))
      WILKS = SSW / SST
      SOMEGA = (SSB - ((DF1*SSW) / DF2)) / (SST + (SSW/DF2))
      RSQ = 1.0 - WILKS
      YSD(1) = SQRT(YSD(1))
      YSD(2) = SQRT(YSD(2))
      YSDP = SQRT(YSDP)
      MESUSD = MESUSD - 1
      WRITE(LPOUTC,2800) MESUSD
2800 FORMAT('1 The following',I4,' measures were used to compute Y.')
      IKK = 1
        DO 340 I = 1,NMEAS
        KK = JGRP2 + IKK
        READ (TEMP1C'KK) XBAR2,SD2,TUK,T,RANG,VAR2
        READ (TEMP1C'KK) KNT2,MIN2,MAX2,MING,MAXG
        IF (TUK .LT. LVLS) GO TO 330
        IPNT = DSL + MCL + PNAME(I) + 1
C---    READ(SSAC'IPNT,ERR = 350) STITLE
        WRITE(LPOUTC,2900) KNT2,(STITLE(NN),NN = 1,STLNG)
2900    FORMAT(' Measure ',I4,': '<STLNG>A1)
330    IKK = 2*I + 1
340    CONTINUE
C
      WRITE(LPOUTC,3000)
3000 FORMAT('0Statistics on new variable:')
      WRITE(LPOUTC,3100)
3100 FORMAT(10X,' MEAN ',20X,' STD.DEV. ')
      WRITE(LPOUTC,3200) YMEAN(1),YSD(1)
3200 FORMAT(' Grp1  ',F10.5,20X,F10.5)
      WRITE(LPOUTC,3300) YMEAN(2),YSD(2)
3300 FORMAT(' Grp2  ',F10.5,20X,F10.5)
      WRITE(LPOUTC,3400) YMEANP,YSDP
3400 FORMAT(' Pooled',F10.5,20X,F10.5)
      WRITE(LPOUTC,3500) F1,F2
3500 FORMAT(/'0 F1 :',G13.6/'  F2 :',G13.6)
      WRITE(LPOUTC,3600) SSB,SSW,SST,WILKS,SOMEGA,RSQ
3600 FORMAT(//'0 SSB  : ',G13.6/'  SSW  : ',G13.6/'  SST  : ',
     1 G13.6///'0 Wilkes-Lambda : ',G13.6/'  S-Omega       : '
     2 ,G13.6/'  R-Squared     : ',G13.6)
      GO TO 360
350 WRITE(TTOUTC,3700) STITLE
3700 FORMAT(' SSATUK - Error while attempting to read measure
     1  name ',<STLNG>A1)
360 RETURN
      END
```

PROGRAM C-2. MANOVA.

```
C     MULTIVARIATE ANALYSIS OF VARIANCE
C     SEE PAGE 238, COOLEY AND LOHNES
C
C     THIS PROGRAM COMPUTES MANOVA TESTS OF H1 (EQUALITY OF DISPERSION
C    AND H2 (EQUALITY OF CENTROIDS), UNIVARIATE F-RATIOS FOR MEANS,
C    SELECTED SAMPLE STATISTICS, AND THE W(POOLED WITHIN-GROUP SSCP.
C    AND T(TOTAL SAMPLE SSCP) MATRICES REQUIRED FOR THE DISCRIMIANT
C    ANALYSIS PROGRAM. THESE MATRICES ARE PUNCHED IN UPPER-TRIANGULAR
C    FORM. THE PROGRAM WILL PROCESS UP TO 20 VARIABLES AND ANY NUMBER
C    OF GROUPS.
C
C     M = NO OF VARIABLES
C     KG = NO OF GROUPS
C     NG = NO OF SUBJECTS , PRECEDING EACH GROUP DATA
C
C     RR - T MATRIX
C     A  - A MATRIX
C     B  - W MATRIX
C     C  - D INVERSE
C     GM - GROUP MEANS, LAST ROW IS THE GRAND MEAN
C
C     SUBROUTINE MATINV IS REQUIRED.
C
 SUBROUTINE MANOVA (M,KG,IO)
C
 INCLUDE 'SALMAI.DEF'
 REAL T(25), AA(25,25), BB(25,25)
C
 WRITE (6,2)
C
2 FORMAT ('1MANOVA.')
 EM=M
 EKG=KG
 EK=KG
 WRITE (6,6) M,KG
6 FORMAT ('0ANALYSIS FOR ',I3,' VARIABLES AND ',I4,' GROUPS')
 WRITE (6,9)
9 FORMAT (1X, 25('-'))
C
C  INITIALIZE VARIABLES
C
 DO 7 J=1,M
 T(J)=0.0
 DO 7 K=1,M
 B(J,K)=0.0
7 C(J,K)=0.0
 H1LOGS=0.0
C
 GA1S=0.0
 FA1S=0.0
 N=0
```

```
C
 DO 19 IG=1,KG
 READ (4) NG
 ENG=NG
 N=N+NG
 WRITE(6,9)
 WRITE(6,10) IG,NG
10 FORMAT (' GROUP ',I3,' NG = ', I6)
 DO 11 J=1,M
 U(J)=0.0
 DO 11 K=1,M
11 A(J,K)=0.0
C
C  CALCULATE MEAN AND S.D.
C
 DO 12 NS=1,NG
 READ(4) (V(J),J=1,M)
 DO 12 J=1,M
 U(J)=U(J)+V(J)
 T(J)=T(J)+V(J)
 DO 12 K=1,M
 A(J,K)=A(J,K)+V(J)*V(K)
12 C(J,K)=C(J,K)+V(J)*V(K)
C
 DO 13 J=1,M
 DO 13 K=1,M
 A(J,K)=A(J,K)-U(J)*U(K)/ENG
 B(J,K)=B(J,K)+A(J,K)
13 A(J,K)=A(J,K)/(ENG-1.)
C
 DO 14 J=1,M
 U(J)=U(J)/ENG
 GM(IG,J) = U(J)
14 W(J)=SQRT (A(J,J))
C
 WRITE(6,15)IG
15 FORMAT (' MEANS FOR GROUP ',I4)
 WRITE (6,16) (U(J),J=1,M)
16 FORMAT (1H0, 10(3X, F7.2))
 WRITE (6,17)
17 FORMAT (21HOSTANDARD DEVIATIONS   )
 WRITE (6,16) (W(J),J=1,M)
C
C  CALCULATE DISPERSION DETERMINANT
C
 CALL MATINV (A,M,DET)
 WRITE (6,18) DET
18 FORMAT ('ODISPERSION DETERMINANT = ',F8.3)
 H1LOGS = H1LOGS+((ENG-1.0)*ALOG (DET))
 FA1S=FA1S+(1.0/(ENG-1.0))
 GA1S=GA1S+(1.0/((ENG-1.0)**2))
19 WRITE(6,9)
C
C  CALCULATE MEANS FOR TOTAL SAMPLE AND S.D.
C  FOR POOLED-SAMPLES
```

```
C
 EN=N
 DO 20 J=1,M
 DO 20 K=1,M
 A(J,K)=C(J,K)-T(J)*T(K)/EN
20 C(J,K)=B(J,K)/(EN-EKG)
C
 DO 21 J=1,M
 T(J)=T(J)/EN
 GM(KG+1,J) = T(J)
21 U(J)=SQRT(C(J,J))
C
 WRITE(6,22)
22 FORMAT ('0MEANS FOR TOTAL SAMPLE')
 WRITE (6,16) ( T(J),J=1,M)
 KGT=KG+1
 WRITE(6,23)
23 FORMAT ('0POOLED-SAMPLES STANDARD DEVIATIONS')
 WRITE (6,16) (U(J),J=1,M)
C
C  SAVE T MATRIX
C
 DO 35 J=1,M
 DO 35 K=1,M
 RR(J,K)=A(J,K)
 AA(J,K)=A(J,K)
 BB(J,K)=B(J,K)
35 A(J,K)=A(J,K)-B(J,K)
C
C     A IS NOW THE A(AMONG-GROUPS SSCP) MATRIX. B IS NOW THE W(WITHIN
C    GROUPS SSCP) MATRIX. C IS NOW THE POOLED-GROUPS DISPERSION EST.
C
C  CALCULATE EQUALITY OF DISPERSION
C
 CALL MATINV(C,M,DET)
 WRITE(6,9)
 H1LOG=(EN-EK)*ALOG (DET)
 XMM=H1LOG-H1LOGS
 F1=.5*(EK-1.0)*EM*(EM+1.0)
 A1A=(FA1S-(1.0/(EN-EK)))*((2.0*(EM*EM))+ (3.0*EM)-1.0)
 A1=A1A/(6.*(EK-1.0)*(EM+1.0))
 A2=(GA1S-(1./(EN-EK)**2))*((EM-1.0)*(EM+2.0))
 2 /(6.0*(EK-1.0))
 DIF=A2-A1*A1
 IF(DIF) 24,24,25
24 F2=(F1+2.0)/(A1*A1-A2)
 B1=F2/(1.0-A1+(2.0/F2))
 F=(F2*XMM)/(F1*(B1-XMM))
 GO TO 45
25 F2=(F1+2.0)/DIF
 B1=F1/(1.0-A1-(F1/F2))
 F=XMM/B1
45 NDF1 = 9999
 IF (F1 .LE. 9999.) NDF1 = F1
 NDF2 = 9999
```

```
 IF (F2 .LE. 9999.) NDF2 = F2
 WRITE(6,26) XMM,F
26 FORMAT('OFOR TEST OF H1 (EQUALITY OF DISPERSIONS), M = ',F10.3,
 2 ' AND F = ', F10.3)
 WRITE (6,27) NDF1, NDF2
27 FORMAT ('OFOR F, NDF1 = ',I3,' AND NDF2 = ',I6)
 PROBF=PRBF(F1,F2,F)
 WRITE (6,900) PROBF
900 FORMAT (' DISPERSION PROB. = ', F8.3)
 WRITE(6,9)
C
C  CALCULATE UNIVARIATE F-RATIOS
C
 N1=EKG-1.0
 N2=EN-EKG
 AN1=N1
 AN2=N2
 WRITE(6,9)
 WRITE(6,40) N1, N2
40 FORMAT('OUNIVARIATE F-RATIOS, WITH NDF1 = ',I3,' AND NDF2 = ',I6)
 WRITE (6,9)
 WRITE (6,41)
41 FORMAT ( 'OVARIABLE AMONG MEAN SQ     WITHIN MEAN SQ F
 1-RATIO ETA SQ    PROB'/)
 DO 42 J=1,M
 ETASQ=A(J,J)/(A(J,J)+B(J,J))
 AMS=A(J,J)/(EKG-1.0)
 WMS=B(J,J)/(EN-EKG)
 F=AMS/WMS
 PROBF=PRBF(AN1,AN2,F)
42 WRITE(6,43)J,AMS,WMS,F,ETASQ,PROBF
43 FORMAT(3X,I3,5XF11.2,11XF11.2,10X,F7.2,8X,F5.4,F8.3)
 WRITE (6,9)
C
 CALL MATINV(BB,M,DETW)
 CALL MATINV (AA,M,DETT)
C
C     DETW IS DETERMINANT OF POOLED-SAMPLES DEVIATION SSCP MATRIX,W.
C     DETT IS DETERMINANT OF TOTAL SAMPLE DEVIATION SSCP MATRIX,T.
C
C  CALCULATE WILKS LAMBDA AND GENERALIZED CORRELATION RATIO
C
 XL=DETW/DETT
 YL=1.0-XL
 WRITE(6,46)XL,YL
 46   FORMAT (16HOWILKS LAMBDA =  F7.4,47H   GENERALIZED CORRELATION RAT
     2IO, ETA SQUARE =  ,F5.4)
C
C  CALCULATE OVERALL DISCRIMINATION
C
 IF(M-2) 47,47,49
 47    IF(KG-3) 48,48,49
C
C     FOR SPECIAL CASES SEE C AND L PAGE 228
C     FOR SMALL NUMBERS OF GROUPS (LT 3), MORE THAN 2 MEASURES
```

```
C
48 KGM1 = KG - 1
 EKGM1 = KGM1
 YL = XL **(1./EKGM1)
 F1 = KGM1*M
 F2 = KGM1*(N-M-KGM1)
 GO TO 50
49 SL = SQRT (((EM * EM) * ((EKG - 1.0)**2) - 4.0)/((EM*EM)+
 2 ((EKG-1.0)**2)-5.0))
 YL=XL**(1.0/SL)
 PL=(EN-1.0)-((EM+EKG)/2.0)
 QL=-((EM*(EKG-1.0))-2.0)/4.0
 RL=(EM*(EKG-1.0))/2.0
 F1=2.0*RL
 F2 = (PL*SL) + (2.0*QL)
50 N1 = 9999
 IF (F1 .LE. 9999.) N1 = F1
 N2 = 9999
 IF (F2 .LE. 9999.) N2 = F2
 F=((1.0-YL)/YL)*(F2/F1)
 WRITE(6,51) F
51 FORMAT('0F-RATIO FOR H2, OVERALL DISCRIMINATION, = ',F9.2)
 WRITE(6,52) N1, N2
52 FORMAT ('0NDF1 = ',I3,' AND NDF2 = ',I6)
 PROBF=PRBF(F1,F2,F)
 WRITE (6,100) PROBF
100 FORMAT (' OVERALL DISCRIMINATION PROB. = ',F8.5)
 WRITE (6,9)
C
 RETURN
 END
```

PROGRAM C-3. MATRIX INVERSION.

```
      SUBROUTINE MATINV(A,M,DET)
C
C GAUSS REDUCTION INVERSION WITHOUT ROW AND COLUMN INTERCHANGES
C
C M IS THE ORDER OF THE SQUARE MATRIX,A
C A-INVERSE IS RETURNED IN A
C DETERMINANT IS RETURNED IN DET
C
      DIMENSION IPIVOT(25),A(25,25),B(25),INDEX(25,2),PIVOT(25)
C
      DET=1.0
      DO 1 J=1,M
      PVT=A(J,J)
      DET=DET*PVT
      A(J,J)=1.0
      DO 2 K=1,M
C  DIVIDE THE PIVOT ROW BY THE PIVOT ELEMENT
 2    A(J,K)=A(J,K)/PVT
      DO 1 K=1,M
C  REDUCE THE NON-PIVOT ROWS
      IF(K-J) 3,1,3
 3    T=A(K,J)
      A(K,J)=0.0
      DO 4 L=1,M
 4    A(K,L)=A(K,L)-A(J,L)*T
 1    CONTINUE
      IF(DET.LT.0.) DET=ABS(DET)
      RETURN
      END
```

PROGRAM C-4. "RIDGE" ADJUSTED MULTIVARIATE DISCRIMINANT ANALYSIS.

```
C      MULTIPLE GROUP DISCRIMINANT ANALYSIS
C
C      THIS PROGRAM COMPUTES DISCRIMINANT FUNCTIONS, THEIR CANONICAL
C     CORRELATIONS WITH GROUP MEMBERSHIP DUMMY VARIATES, F-RATIOS FOR
C     THESE, AND CENTROIDS OF GROUPS IN THE STANDARDIZED DISCRIMINANT
C     FUNCTIONS SPACE. COEFFICIENTS FOR COMPUTING STANDARDIZED
C     DISCRIMINANT FUNCTIONS SCORES FROM DEVIATION TEST SCORES ARE
C      SEE PAGE 258, COOLEY AND LOHNES 1971
C
C      REQUIRED SUBROUTINES ARE DIRNM AND HOW.
C
C      INPUT
C
C      M = NO. OF VARIABLES
C      KG = NO. OF GROUPS
C      N = NO. OF SUBJECTS
C      RR = T-MATRIX, TOTAL SAMPLE DEV. AS OUPUT BY MANOVA
C      B = W-MATRIX, POOLED WITHIN-GROUPS DEV.,FROM MANOVA
C      GM-MATRIX, GROUP MEANS AND GRAND MEANS, FROM MANOVA
C      S = "RIDGE" ADJUSTMENT
C
C      OUTPUT
C
C      XL = WILKS LAMBDA
C      V = CANONICAL R
C      Y = CHI-SQUARE
C      A = FACTOR PATTERN FOR DISCRIMINANT FUNCTIONS
C
C
       SUBROUTINE DISCRM (M,KG,N,IO)
C
          DIMENSION A(25,25), B(25,25), C(25,25), T(25), U(25), V(25),
         2 W(25), X(25), Y(25), Z(25), D(25,25), GM(25,25)
C
          COMMON /MAT/C,A,B,D,GM,V,Y,XL
          COMMON /SCR/   U,W,X,Z
          COMMON /S/S
C
 REAL T(25)
C
 KC = 0
1 WRITE (6,2)
2 FORMAT('1MULTIPLE GROUP DISCRIMINANT ANALYSIS')
 DO 8 J=1,M
 DO 8 K=J,M
 RR(K,J)=RR(J,K)
8 B(K,J)=B(J,K)
C
 DO 15 J=1,M
 DO 15 K=1,M
15 A(J,K)=RR(J,K)-B(J,K)
```

```
      DO 500 I=1,M
500   B(I,I)=B(I,I)+S
C
C     A NOW CONTAINS THE A MATRIX (AMONG-GROUPS DEVIATION SSCP MATRIX).
C     B CONTAINS THE W MATRIX (WITHIN-GROUPS DEVIATION SSCP MATRIX)
C (INCLUDING "RIDGE" ADJUSTMENT).
C
      IF (M-KG) 10,11,11
10    MD=M
      GO TO 12
11    MD=KG-1
C
12    CALL DIRNM (A,M,B,D,T,MD)
C
C     ROOTS OF W INVERSE*A ARE IN T AND COLUMN EIGENVECTORS ARE IN D.
C
      EM=M
      EKG=KG
      EN=N
      EKC=KC
      XL=1.0
      TRACE =0.0
      DO 13 J=1,MD
      U(J)=T(J)/(1.0+T(J))
      V(J)=SQRT (U(J))
      W(J)=1.0/(1.0+T(J))
      XL=XL*W(J)
D     WRITE (5,9100) T(J),U(J),V(J),W(J),XL,TRACE
D9100 FORMAT (' T = ',E16.7,' U = ',E16.7,' V = ',E16.7/
D    1 ' W = ',E16.7,' XL= ',E16.7,' TR= ',E16.7)
13    TRACE=TRACE+T(J)
      CR=TRACE/(MD+TRACE)
      ES=SQRT((CR*CR)/(1-CR*CR))
      OUTPUT ES,CR
D     WRITE (5,9000) ES,CR
D9000 FORMAT (' ES = ',E16.7,' CR = ',E16.7)
C
      DO 14 J=1,MD
14    Z(J)=100.0*(T(J)/TRACE)
C
      IF (M-2) 16,16,17
16    IF (KG-3) 18,18,17
C
C     FOR SPECIAL CASES SEE C AND L PAGE 228
C     FOR SMALL NUMBERS OF GROUPS (LT 3), MORE THAN 2 MEASURES
C
18    KGM1 = KG - 1
      EKGM1 = KGM1
      YL = XL **(1./EKGM1)
      F1 = KGM1*M
      F2 = KGM1*(N-M-KGM1)
      GO TO 19
17    SL=SQRT(((EM*EM)*((EKG-1.0)**2)-4.0)/((EM*EM)+
     2 ((EKG-1.0)**2)-5.0))
      YL=XL**(1.0/SL)
```

```
      PL=(EN-1.0-EKC)-((EM+EKG)/2.0)
      QL=-((EM*(EKG-1.0))-2.0)/4.0
      RL=(EM*(EKG-1.0))/2.0
      F1=2.0*RL
      F2=(PL*SL)+(2.0*QL)
   19 N1 = 9999
      IF (F1 .LE. 9999.) N1 = F1
      N2 = 9999
      IF (F2 .LE. 9999.) N2 = F2
      F=((1.0-YL)/YL)*(F2/F1)
      YL=1.0-XL
      WRITE  (6,201)XL,YL
  201 FORMAT('OWILKS LAMBDA = ',F7.4,' GENERALIZED  CORRELATION RATIO,',
     2 ' ETA SQUARED = ',F7.4)
      WRITE(6,20)F
   20 FORMAT('OF-RATIO FOR H2,OVERALL DISCRIMINATION = ',E16.7)
      WRITE (6,21) N1, N2
   21 FORMAT('ONDF1  = ',I3, ' AND NDF2 = ',I6)
      PROBF=PRBF(F1,F2,F)
      WRITE (6,211) PROBF
  211 FORMAT (' OVERALL DISCRIM. PROB. = ',F8.3)
      J=MD
      X(J+1)=1.0
   22 X(J)=X(J+1)*W(J)
      J=J-1
      IF(J) 23,23,  22
   23 DO 24 J=1,MD
   24 Y(J)=-PL*ALOG(X(J))
C
      WRITE(6,25)
   25 FORMAT('OCHI-SQUARE TESTS WITH SUCCESSIVE ROOTS REMOVED')
      WRITE (6,261)
  261 FORMAT('0',19X,'(ETA) (ETA SQUARE )')
      WRITE (6,26)
   26 FORMAT (' ROOTS REMOVED  CANONICAL R    R SQUARED    EIGEN',
     1 'VALUE   CHI-SQUARE      N D F    LAMBDA  PERCENT TRACE'/)
      DO 27 J=1,MD
      JT=J-1
      NDF=(M-JT)*(KG-JT-1.0)
   27 WRITE(6,28) JT,V(J),U(J),T(J),Y(J),NDF, X(J),Z(J)
C
   28 FORMAT(6X,I4,9X,2(F6.3,8X),E16.7,5X,F10.2,4X,I5,2X,F9.2,F8.2)
C
C     D = COEFFICIENTS VECTORS
C
      DO 29 J=1,MD
      DO 29 K=1,M
      A(J,K)=0.0
      DO 29 L=1,M
   29   A(J,K)=A(J,K)+D(L,J)*(RR(L,K)/(EN-1.0))
C
      DO 30 J=1,MD
      DO 30 K=1,MD
      B(J,K)=0.0
      DO 30 L=1,M
```

```
30 B(J,K)=B(J,K)+A(J,L)*D(L,K)
C
 DO 31 J=1,M
 DO 31 K=1,MD
31 D(J,K)=D(J,K)*(1.0/SQRT(B(K,K)))
C
 WRITE (6,32)
32 FORMAT('OROW COEFFICIENTS VECTORS')
 DO 33 J=1,MD
C
33 WRITE(6,49) J, (D(K,J),K=1,M)
49 FORMAT(' D F ',I3,2X,5F8.3/(10X,5F8.3))
 DO 34 J=1,M
34 Z(J)=SQRT(RR(J,J)/(EN-1.0))
C
C     TOTAL SAMPLE STANDARD DEVIATIONS ARE NOW IN Z.
C
 DO 35 J=1,M
 DO 35 K=1,M
35 RR(J,K)=RR(J,K)/(EN*Z(J)*Z(K))
C
C     TOTAL SAMPLE CORRELATION MATRIX IS NOW IN C.
C
 DO 36 J=1,M
 DO 36 K=1,MD
36 B(J,K)=D(J,K)*Z(J)
C
 DO 37 J=1,M
 DO 37 K=1,MD
 A(J,K)=0.0
 DO 37 L=1,M
37 A(J,K)=A(J,K)+RR(J,L)*B(L,K)
C
 WRITE(6,38)
38 FORMAT('OFACTOR PATTERN FOR DISCIMINANT FUNCTIONS')
 DO 39 J=1,M
39 WRITE (6,40) J, (A(J,K),K=1,MD)
C
40 FORMAT (' TEST ',I4, 10(3X,F7.3) / (9X, 10(3X,F7.3)))
 DO 41 J=1,M
 T(J)=0.0
 DO 41 K=1,MD
41 T(J)=T(J)+A(J,K)*A(J,K)
C
 WRITE (6,42) MD
42 FORMAT('OCOMMUNALITIES FOR  ',I5,' DISCRIMINANT FACTORS')
 WRITE (6,43) (J,T(J), J=1,M)
43 FORMAT ('0', 10(2X, I3, F7.3))
C
C     SAVE COMMUNALITIES IN THE V ARRAY
C
 DO 231 J=1,M
231 V(J)=T(J)
C
 DO 44 J=1,MD
```

```
      T(J)=0.0
      DO 44 K=1,M
   44 T(J)=T(J)+A(K,J)*A(K,J)
C
      WRITE (6,45)
   45 FORMAT ('OPERCENTAGE OF TRACE OF R ACCOUNTED FOR BY EACH ROOT')
      DO 46 J=1,MD
   46 T(J)=100.0*(T(J)/EM)
      WRITE (6,43) (J,T(J),J=1,MD)
C
      KGT=KG+1
C
C     READS GROUP MEAN VECTORS AND GRAND MEAN VECTOR INTO COLUMNS OF A.
C     COLUMN KGT CONTAINS THE GRAND MEANS.
C
      DO 47 J=1,KGT
      DO 47 K=1,M
   47 A(J,K) = GM(J,K)
C
      DO 48 J=1,KG
      DO 51 K=1,MD
      T(K)=0.0
      DO 51 L=1,M
   51 T(K)=T(K)+(A(J,L)-A(KGT,L))*D(L,K)
      WRITE (6,50) J,MD
   50 FORMAT ('OCENTROID FOR GROUP  ',I4,' IN ',I4,' DIMENSIONAL DISCRIM',
     2 'INANT SPACE')
   48 WRITE(6,43) (K,T(K),K=1,MD)
C
      RETURN
      END
```

PROGRAM C-5. DIAGONALIZATION OF A REAL NON-SYMMETRIC MATRIX.

(OF THE FORM B-INVERSE*A. CODED BY P.R. LOHNES,U.N.H.)

```
C
C     A,M,B,X,AND XL ARE DUMMY NAMES AND MAY BE CHANGED IN THE
C CALLING STATEMENT.
C
C     THE EIGENVALUES OF B-1*A,AND MATRIX X CONTAINS THE EIGENVECTORS
C     IN ITS COLUMNS.  SUBROUTINE HOW PACKAGE IS REQUIRED.
C
C     LVECT SPECIFIES THE NUMBER OF EIGENVECTORS TO BE RETURNED.
C
      SUBROUTINE DIRNM (A,M,B,X,XL,LVECT)
C
      REAL A(25,25),B(25,25),X(25,25),XL(25),U(25),V(25),W(25),Y(25)
C
      CALL  HOW(M,25,M,B,XL,X,U,V,W,Y)
C
      DO 1 I=1,M
 1    XL(I)=1.0/SQRT(ABS(XL(I)))
      DO 2 I=1,M
      DO 2 J=1,M
 2    B(I,J)=X(I,J)*XL(J)
      DO 3 I=1,M
      DO 3 J=1,M
      X(I,J)=0.0
      DO 3 K=1,M
 3    X(I,J)=X(I,J)+B(K,I)*A(K,J)
      DO 4 I=1,M
      DO 4 J=1,M
      A(I,J)=0.0
 DO 4 K=1,M
4 A(I,J)=A(I,J)+X(I,K)* B(K,J)
C
C     A NOW CONTAINS BPRIME*A*B OF THE NOTES.
C
      CALL  HOW(M,25,LVECT,A,XL,X,U,V,W,Y)
C
      DO 6 I=1,M
      DO 6 J=1,M
      A(I,J)=0.0
      DO 6 K=1,M
 6    A(I,J)=A(I,J)+B(I,K)*X(K,J)
      DO 9 I=1,M
      SUMV=0.0
      DO 7 J=1,M
 7    SUMV=SUMV+(A(J,I)**2)
      DEN=SQRT (SUMV)
      DO 8 J=1,M
 8    X(J,I)=A(J,I)/DEN
 9    CONTINUE
C
C     COLUMNS OF X(I,J) ARE NOW NORMALIZED.
```

```
C
      RETURN
      END
```

PROGRAM C-6. EIGENVALUES AND EIGENVECTORS OF A SYMMETRIC MATRIX.

```
C        BY HOUSEHOLDER,ORTEGA, AND WILKINSON.  ORIGINAL PROGRAM BY
C     DAVID W. MATULA UNDER THE DIRECTION OF WILLIAM MEREDITH,
C      UNIVERSITY OF CALIFORNIA, BERKELEY, 1962
C       (SEE RALSTON AND WILF,VOLUME II(1967))
C      MODIFIED BY P.R. LOHNES, PROJECT TALENT,1966.
C
C      M IS THE ORDER OF THE INPUT MATRIX,R.
C      MD IS THE DIMENSIONED SIZE OF R IN THE MAIN PROGRAM.
C      NV IS THE NUMBER OF EIGENVECTORS TO BE COMPUTED.
C      E IS THE VECTOR IN WHICH THE EIGENVALUES ARE RETURNED.
C      V IS THE MATRIX IN WHICH THE EIGENVECTORS ARE RETURNED.
C         THE EIGENVECTORS ATE STORED AS COLUMNS IN V.
C      A,B,C, AND D ARE WORKSPACE VECTORS.
C
       SUBROUTINE HOW (MVAR,MDIM, NVECT,R,E,V,A,B,C,D)
C
       DIMENSION R(625),E(25),V(625),A(25),B(25),C(25),D(25)
C
       M=MVAR
       MD=MDIM
       NV= NVECT
C
       IF (M-1) 100,97,96
C
 96    M1=M-1
C
C  TRI-DIAGONALIZE THE MATRIX.  (HOUSEHOLDER'S SIMILARITY ORTHOGONAL
C  TRANSFORMATION)
C
       M2=M1*MD+M
       M3=M2-MD
       M4=MD+1
       L=0
       DO 1 I=1, M2,M4
       L=L+1
 1     A(L)=R(I)
       B(1)=0.0
C
       IF(M-2) 13,2,3
C
 3     KK=0
       DO15 K=2,M1
       KL=KK+K
       KU=KK+M
       KJ=K+1
       SUM=0.0
       DO 4 J=KL,KU
 4     SUM=SUM+R(J) **2
       S=SQRT(SUM)
       Z=R(KL)
       B(K)=SIGN(S,-Z)
```

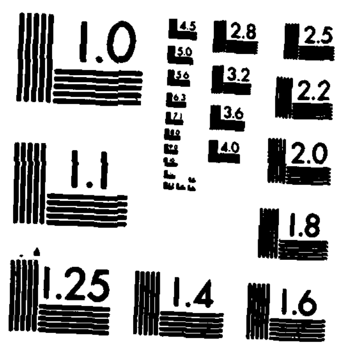
```
      S=1.0/S
      C(K)=SQRT (ABS (Z)*S+1.0)
      X=SIGN (S/C(K),Z)
      R(KL)=C(K)
      DO 5 I=KJ,M
      JJ=I+KK
      C(I)=X*R(JJ)
 5    R(JJ)=C(I)
      DO 8 J=K,M
      JJ=J+1
      D(J) =0.0
      L=KK+J
      DO 6 I=K,J
      L=L+MD
 6    D(J)=D(J) +R(L) *C(I)
C
      IF (JJ-M) 7,7,9
C
 7    DO8 I=JJ,M
      L=L+1
 8    D(J)=D(J) +R(L)*C(I)
 9    X=0.0
      DO 10 J=K,M
 10   X=X+C(J) *D(J)
      X=.50*X
      DO11 I=K,M
 11    D(I)=X*C(I)-D(I)
      LL=KK
      KK=KK+MD
      DO 15 I=K,M
      LL=LL+MD
      DO15 J=I,M
      L=LL+J
 15    R(L)=R(L)+D(I)*C(J)+D(J)*C(I)
C
C
      L=1
      DO 12 I=1,M
      X=A(I)
      A(I)=R(L)
      R(L)=X
 12    L=L+M4
 2    B(M)=R(M3)
C
C  COMPUTE EIGENVALUES.  (ORTEGA'S METHOD OF STURM SEQUENCES)
C
 13    BD=ABS (A(1))
      DO14 I=2,M
 14    BD= AMAX1(BD,ABS (A(I))+B(I)**2)
      BD=BD+1.0
      DO 16 I=1,M
      A(I)=A(I)/BD
      B(I)=B(I)/BD
      D(I)=1.0
 16    E(I)=-1.0
```

```
      DO 37 K=1,M
C
 17    IF((D(K)-E(K))/AMAX1(ABS (D(K)),ABS (E(K)),1.0E-9)-1.0E-6)
     2  37,37,18
C
 18    X=(D(K)+E(K))*.50
      IS2=1
 S2 = 1.
      C(1)=A(1)-X
C
      IF(C(1)) 19,20,20
C
 19    IS1=-1
 S1 = -1.
      N=0
      GO TO 21
C
 20   IS1=1
 S1 = 1.
      N=1
 21    DO 31 I = 2,M
C
      IF(B(I)) 22,26,22
 22    IF(B(I-1)) 23,27,23
 23    IF(ABS(C(I-1))+ABS(C(I-2))-1.0E-15) 24,25,25
C
24    C(I-1)=C(I-1)*1.0E15
      C(I-2)=C(I-2)*1.0E15
 25    C(I)=(A(I)-X)*C(I-1)-B(I)**2*C(I-2)
      GO TO 28
C
 26    C(I)=(A(I)-X)*SIGN (1.,S1)
      GO TO 28
C
 27    C(I)=(A(I)-X)*C(I-1)-SIGN (B(I)**2,S2)
 28    S2=S1
C
      IF(C(I)) 29,30,29
C
 29    S1=SIGN (S1,C(I))
C
      IF(S2+S1) 30,31,30
C
 30    N=N+1
 31    CONTINUE
      N=M-N
C
      IF (N-K) 34,32,32
C
 32    DO 33 J=K,N
 33    D(J)=X
 34    N=N+1
C
      IF(M-N) 17,35,35
C
```

AD-A159 055 PERFORMANCE MEASUREMENT GUIDELINES FOR RESEARCH(U) 3/3
VREULS RESEARCH CORP THOUSAND OAKS CA D VREULS ET AL.
13 JUN 85 AFOSR-TR-85-0642 F49620-80-C-0058

UNCLASSIFIED F/G 5/9 NL

END
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

```
 35    DO 36 J=N,M
C
      IF(X-E(J)) 17,17,36
C
36    E(J)=X
      GO TO 17
C
37 CONTINUE  !CSB 01/29/79
C
C
      DO 38 I=1,M
      A(I)=A(I)*BD
       B(I)=B(I)*BD
 38    C(I)=(D(I)+E(I))*BD*.50
      M1=M
      K=1
 39    I=1
 40    DO 43 J=1,M1
C
      IF (I-J) 41,43,41
 41    IF(C(I)-C(J)) 43,43,42
C
 42    I=J
      GO TO 40
C
 43    CONTINUE
      E(K)=C(I)
      K=K+1
      M1=M1-1
C
      IF(I-M1-1) 44,46,46
C
 44    DO 45 M2=1,M1
 45    C(M2)=C(M2+1)
 46    IF(M1-1) 47,47,39
C
 47    E(K) = C(1)
C
      IF (ISIGN (1,NV)) 79,76,76
C
 76    DO 77 I=1,M
 77    C(I)=E(I)
      J=M
      DO 78 I=1,M
      E(I)=C(J)
 78    J=J-1
79    CONTINUE
C
C  DECIDE WHETHER TO COMPUTE EIGENVECTORS,AND IF SO,HOW MANY
C  (WILKINSON'S PROCEDURE)
C
      IF(NV) 48,99,48
C
 48    KX= IABS(NV)
      J=1
```

```
      DO 98 INV=1,KX
      X=A(1)-E(INV)
      Y=B(2)
      M1=M-1
      DO 54 I=1,M1
      IJ=J+I-1
C
      IF(ABS(X)-ABS(B(I+1))) 49,51,53
C
 49    C(I)=B(I+1)
      D(I)=A(I+1) -E(INV)
      V(IJ)=B(I+2)
      Z=-X/C(I)
      X=Z*D(I)+Y
C
      IF (M1-I) 50,54,50
C
 50    Y=Z* V(IJ)
      GO TO 54
C
 51    IF (X) 53,52,53
C
 52    X=1.0E-10
 53    C(I)=X
      D(I)=Y
      V(IJ)=0.0
      X=A(I+1)-(B(I+1)/X*Y+E(INV))
      Y=B(I+2)
 54    CONTINUE
      MJ=M+J-1
C
      IF(X) 56,60,56
C
 56    V(MJ)=1.0/X
 57    I=M1
      IJ=J+I-1
      V(IJ)=(1.0-D(I)*V(MJ))/C(I)
      X=V(MJ)**2+V(IJ)**2
 58    I=I-1
      IJ =J+I-1
C
      IF(I) 59,61,59
C
 59    V(IJ)=(1.0-(D(I)*V(IJ+1)+V(IJ)*V(IJ+2)))/C(I)
      X=X+V(IJ)**2
      GO TO 58
C
 60    V(MJ)=1.0E10
      GO TO 57
C
 61    X=SQRT(X)
      DO 62 I=1,M
      IJ=J+I-1
 62    V(IJ)=V(IJ)/X
      J1=M1*MD-MD
```

```
      K=M
      GO TO 66
C
 63     K=K-1
      J1=J1-MD
      Y=0.0
      DO 64 I=K,M
      IJ=J+I-1
      L=J1+I
 64    Y=Y+V(IJ)*R(L)
      DO 65 I=K,M
      IJ=J+I-1
      L=J1+I
 65    V(IJ)=V(IJ)-Y*R(L)
C
 66    IF (J1) 63,67,63
C
 67    NPLUS =0
      NMIN =0
      DO 70 I=1,M
      IJ=J+I-1
C
      IF (V(IJ)) 68,69,69
C
 68    NMIN=NMIN+1
      GO TO 70
C
 69    NPLUS=NPLUS+1
 70    CONTINUE
C
      IF (NPLUS-NMIN) 71,73,73
C
 71    DO 72 I=1,M
      IJ=J+I-1
 72    V(IJ)=-V(IJ)
 73    CONTINUE
 98    J=J+MD
C
C
C     RESTORE THE INPUT MATRIX.
C
 99    MD1=MD+1
      JJ=MD1
      M1=M*MD
      DO 75 I=2,M1,MD1
      K=I
      DO 74 J=JJ,M1,MD
      R(K)=R(J)
 74    K=K+1
 75    JJ=JJ+MD1
      GO TO 100
C
 97    E(1)=R(1)
      V(1)=1.0
C
```

```
100     RETURN
        END
```

PROGRAM C-7. LOWER TRIANGULAR SQUARE ROOT OF A MATRIX.

```
 SUBROUTINE MXSQRT(M)
C
C  M = DIMENSION OF THE REAL, SYMMETRIC, VARIANCE-COVARIANCE MATRIX, D
C
 INCLUDE 'SALMAI.DEF'
C
C  ZERO ALL ENTRIES IN DUMMY MATRIX C
C
 DO 100 I=1,25
 DO 50 J=1,25
 C(I,J)=0.0
50 CONTINUE
100 CONTINUE
C
C  COMPUTE FIRST COLUMN
C
 ASQRT=SQRT(D(1,1))
 DO 101 I=1,M
 C(I,1)=D(I,1)/ASQRT
101 CONTINUE
C
C  COMPUTE SECOND THRU MTH ROWS STARTING WITH SECOND COLUMN OF EACH ROW
C
C  COMPUTE SECOND MAIN DIAGONAL ENTRY, C(2,2),  ALL OTHER MAIN DIAGONAL
C  BE COMPUTED  LATER
C
 C(2,2)=SQRT(D(2,2)-C(2,1)*C(2,1))
C
C  NOW, COMPUTE REMAINDER OF ENTRIES
C
 DO 102 I=3,M
 DO 103 J=2,I-1
 SUM=0.0
C
 DO 104 K=1,J-1
 SUM=C(I,K)*C(J,K) + SUM
104 CONTINUE
C
 C(I,J)=(D(I,J)-SUM)/C(J,J)
103 CONTINUE
C
 SUM=0.0
 DO 105 K=1,I-1
 SUM=C(I,K)*C(I,K) + SUM
105 CONTINUE
C
 C(I,I) = SQRT(D(I,I)-SUM)
102 CONTINUE
C
 DO 200 I=1,M
 DO 200 J=1,M
```

```
200 D(I,J)=C(I,J)
C
C RECONSTRUCT VARIANCE-COVARIANCE MATRIX
C FORM TRANSPOSE OF SQUARE ROOT MATRIX
C
250 DO 300 I=1,M
 DO 300 J=1,M
300 RR(I,J)=C(J,I)
C
C FORM THE PRODUCT(SQUARE)
C
 DO 400 J=1,M
 DO 400 I=1,M
 SUM=0.0
 DO 500 K=1,M
500 SUM=SUM+D(I,K)*RR(K,J)
 C(I,J)=SUM
400 CONTINUE
 RETURN
 END
```

PROGRAM C-8. CREATE SIMULATION DATA.

```
      SUBROUTINE SDATA(M,NG,NS,IUNIT)
C
C   THIS ROUTINE CREATES SIM. DATA FROM REAL
C
C
         DIMENSION X(25),SX(25),SS(25,25),SSD(25,25),D(25,25)
         DIMENSION R(25,25),XM(25),SD(25),VAR(200),IVAR(25),Y(25)
         COMMON /SCR/ X,SX,SS,SSD,D,XM,SD
 REAL VAR(200), XM(25)
 REWIND(IUNIT)
 REWIND(2)
 DO 1000 ISUP=1,2
 DO 12 I=1,M
12 XM(I)=0.0
C
C  READ ACTUAL DATA
C
C  READ NUMBER OF OBSERVATIONS
C
 READ (IUNIT) NG
 REWIND 3
 WRITE (3) NG
 DO 10 I=1,NG
 READ(IUNIT) (X(L),L=1,M)
 DO 11 L=1,M
11 XM(L)=XM(L)+X(L)
10 WRITE (3) (X(L),L=1,M)
 DO 13 I=1,M
13 XM(L)=XM(L)/NG
 REWIND (3)
C
C  FIND CORREL. MAT. AND VAR. COV. MAT.
C
 CALL CORREL(M,R,1,0)
C
C  DETERMINE SQUARE ROOT MAT.
C
 CALL MXSQRT(M)
C
C  PRODUCE RANDOM NORMAL DIST.
C
 IF(NS.EQ.0) NS=NG
 IP=12345
 IPOINT=1
 REWIND (3)
 WRITE (3) NS
 DO 60 K=1,NS
 DO 50 I=1,M,2
 CALL RAND(IP,Y(1))
 CALL RAND(IP,Y(2))
 VAR(I)=SQRT(-2*ALOG(Y(1)))*COS(6.283185*Y(2))
```

```
50 VAR(I+1)=SQRT(-2*ALOG(Y(1)))*SIN(6.283185*Y(2))
C
C  PRODUCE NEW RANDOM VAR.
C
 DO 71 J=1,M
 X(J)=0.0
 DO 70 JJ=1,J
70 X(J)=X(J)+VAR(J)*D(J,JJ)
71 X(J)=X(J)+XM(J)
C
C  WRITE NEW DATA OUT
C
 WRITE (2) (X(L),L=1,M)
 WRITE (3) (X(L),L=1,M)
60 CONTINUE
 REWIND (3)
1000 CONTINUE
 RETURN
 END
```

PROBRAM C-9. COMPUTE PROBABILITY OF F-RATIO.

```
 SUBROUTINE PRBF(DA,DB,FR)
C
C COMPUTE 95% F-RATIO PROBABILITY BY APPROXIMATING DENSITY FUNCTION
C INTEGRAL.  EXCERPT FROM VANDERBILT STAT PACKAGE.
C
C
C WHERE:
C  DA = NUMERATOR DEGREES OF FREEDOM
C  DB = DENOMINATOR     "  OF   "
C  FR = F-RATIO
C
 DATA C0,C1,C2,C3,C4 /1.0, 0.196854, 0.115194, 0.000344, 0.019527/
C
 PRBF = 1.
 IF (DA .LE. 0.  .OR.  DB .LE. 0.) RETURN
 IF (FR .LE. 0.)     RETURN
C
 A = DA
 B = DB
 F = FR
 IF (F .GT. 0.) GO TO 10
 A = DB
 B = DA
 F = 1./FR
10 AA = 2./(9.*A)
 BB = 2./(9.*B)
 FC = F**0.3333333
C
 Z = ABS ( (1.-BB)*FC - (1.-AA))/
 1 SQRT (BB*FC**2 + AA)
 IF (B .LT. 4.) Z = Z*(1. + 0.08*Z**4/B**3)
C
 PRBF = 0.5/
 1 (C0+Z*(C1+Z*(C2+Z*(C3+Z*C4))))
C
 IF (FR .LT. 1.) PRBF = 1.-PRBF
 RETURN
 END
```

PROGRAM C-10. COMPUTE POWER OF F-TEST.

```
 SUBROUTINE FPOWR(V1,V2,F,PAR)
C
C WOODWARD AND OVERALL (1976)
C COMPUTES THE POWER OF THE F-TEST
C
C V1 - NDF1
C V2 - NDF2
C F  - VALUE OF THE CENTRAL F-DISTRIBUTION
C PAR- NON-CENTRALITY PARAMETER
C
 A=((V1*F)/(V1+PAR))**0.333333
 B=(1.0-(2.0/(9.0*V2)))
 C=1.0-(2.0*(V1+2.0*PAR))/(9.0*((V1+PAR)**2.0))
 D=(2.0(V1+2.0*PAR))/(9.0*((V1+PAR)**2))
 E=(2.0/(9.0*V2))*(((V1*F)/(V1+PAR))**0.666666)
 D=SQRT(D+E)
 Z=(A*B-C)/D
 C1=.196854
 C2=.115194
 C3=.000344
 C4=.019527
 IF(Z) 201,200,200
200 P=0.5/((1.0+C1*Z+C2*(Z+Z)+C3*(Z*Z*Z)-C4*(Z*Z*Z*Z))**4)
 GO TO 99
201 Z=ABS(Z)
 P=0.5/((1.0+C1*Z+C2*(Z*Z)+C3*(Z*Z*Z)-C4*(Z*Z*Z*Z))**4)
 P=1.0-P
99 WRITE(OUTC,1) P
1 FORMAT(' POWER OF F-TEST='F10.5)
C
 RETURN
 END
```

PROGRAM C-11. REPEATED MEASURES ADJUSTMENT.

```
      SUBROUTINE SUTR(NG,MAN11)
C
C SUBTARCT SUBJECT AND TRIAL EFFECTS FROM RAW DATA.
C
      COMMON TABLE(12,9,6,12)
      COMMON X(30),SUMD(30,2),SUMDS(30,2,12),SUMDT(30,2,12)
      COMMON /JUNK/NAME,RNAME,NCHECK,M,MHOLD,MSTART
      COMMON /REP/ISUBJR,NAG
C
      INTEGER NAME(40),RNAME(40),NCHECK(40),MSTAT(40),ISUBJR(12),NTR(12),
     1  NAG(9),TABLE
C
      REWIND(10)
      REWIND(11)
C
C INITIALIZE SUMS
C
      DO 10 L=1,30
      DO 10 I=1,2
      SUMD(L,I)=0.
      DO 10 K=1,12
      SUMDS(L,I,K)=0.
10    SUMDT(L,I,K)=0.
C
C GET SUMS
C
      DO 30 I=1,2
      READ(10) N
      WRITE(11) N
      DO 20 J=1,12
20    NTR(J)=0
      DO 30 J=1,12
      K2=0
      DO 30 K=1,ISUBJR(J)
      K2=K2+1
      READ(10) (X(L),L=1,M)
      WRITE(11) (X(L),L=1,M)
7     IF(TABLE(J,NAG(1),MAN11,K2).NE.0) GO TO 9
      IF(TABLE(J,NAG(2),MAN11,K2).EQ.0) GO TO 11
9     K2=K2+1
      GO TO 7
11    NTR(K2)=NTR(K2)+1
      DO 30 L=1,M
      Y=X(L)
      SUMD(L,I)=SUMD(L,I)+Y
      SUMDS(L,I,J)=SUMDS(L,I,J)+Y
      SUMDT(L,I,K2)=SUMDT(L,I,K2)+Y
30    CONTINUE
C
C GET MEANS
C
```

```
      DO 40 L=1,M
      DO 40 I=1,2
      SUMD(L,I)=SUMD(L,I)/N
      DO 50 J=1,12
   50 SUMDS(L,I,J)=SUMDS(L,I,J)/ISUBJR(J)
      DO 40 K=1,12
   40 SUMDT(L,I,K)=SUMDT(L,I,K)/NTR(K)
C
      DO 120 L=1,5
      WRITE(8,111) L,SUMD(L,1),SUMD(L,2)
      DO 140 J=1,12
  140 WRITE(8,111) J,SUMDS(L,1,J),SUMDS(L,2,J)
      DO 120 I=1,2
  120 WRITE(8,113) (SUMDT(L,I,K),K=1,12)
  111 FORMAT(I5,2F8.4)
  113 FORMAT(12F8.4/)
C
C SUBTRACT RESULTS FROM RAW DATA
C
      REWIND(10)
      REWIND(11)
      DO 60 I=1,2
      READ(11) N
      WRITE(10) N
      DO 60 J=1,12
      K2=0
      DO 60 K=1,ISUBJR(J)
      K2=K2+1
   13 IF(TABLE(J,NAG(1),MAN11,K2).NE.0) GO TO 15
      IF(TABLE(J,NAG(2),MAN11,K2).EQ.0) GO TO 17
   15 K2=K2+1
      GO TO 13
   17 READ(11) (X(L),L=1,M)
      DO 80 L=1,M
   80 X(L)=X(L)-SUMDS(L,I,J)-SUMDT(L,I,K2)+2.*SUMD(L,I)
   60 WRITE(10) (X(L),L=1,M)
C
      REWIND(10)
      REWIND(11)
      RETURN
      END
```

END

FILMED

10-85

DTIC